

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP  
CAMPUS DE JABOTICABAL**

**IDENTIFICAÇÃO DE HAPLÓTIPOS HOMOZIGOTOS  
PORTADORES DE POSSÍVEIS MUTAÇÕES LETAIS E  
SEMILETAIS RECESSIVAS EM POPULAÇÕES DE BOVINOS  
GIR E GIROLANDO**

**Bárbara Carolina Gomes Ribeiro**

Matemática e Engenharia Civil

**UNIVERSIDADE ESTADUAL PAULISTA  
CAMPUS DE JABOTICABAL**

**IDENTIFICAÇÃO DE HAPLÓTIPOS HOMOZIGOTOS  
PORTADORES DE POSSÍVEIS MUTAÇÕES LETAIS E  
SEMILETAIS RECESSIVAS EM POPULAÇÕES DE BOVINOS  
GIR E GIROLANDO**

**Bárbara Carolina Gomes Ribeiro**

Orientador: **Prof. Dr. Fernando Sebastián Baldi Rey**

Coorientador: **Prof. Dr. Marcos Vinicius Gualberto Barbosa da Silva**

Coorientadora: **Profa. Dra. Elisa Peripolli**

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

R484i

Ribeiro, Bárbara Carolina Gomes

Identificação de haplótipos homocigotos portadores de possíveis mutações letais e semiletais recessivas em populações de bovinos Gir e Girolando / Bárbara Carolina Gomes Ribeiro. -- Jaboticabal, 2025  
80 f. : tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP),  
Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientador: Fernando Sebastián Baldi Rey

Coorientador: Marcos Vinícius Gualberto Barbosa Silva

1. Alelos recessivos. 2. Endogamia. 3. Inferência de haplótipos. 4.  
Fertilidade. 5. Indivíduos homocigotos. I. Título.

## **IMPACTO POTENCIAL DESTA PESQUISA**

Esta pesquisa oferece benefícios significativos à sociedade ao aprimorar a eficiência dos programas de melhoramento genético de bovinos das raças Gir e Girolando, contribuindo para a redução de perdas econômicas associadas a alelos recessivos letais e semiletais. Ao identificar e mapear haplótipos letais, bem como loci segregantes entre as raças, esta pesquisa fornece ferramentas para evitar acasalamentos que resultem em progênes inviáveis, promovendo maior eficiência reprodutiva e produtiva. Além disso, os resultados podem subsidiar ações nos programas de melhoramento genético, como o Programa Nacional de Melhoramento do Gir Leiteiro (PNMGL) e o Programa de Melhoramento Genético do Girolando (PMGG), impactando diretamente a sustentabilidade econômica e ambiental da bovinocultura leiteira brasileira, setor fundamental para a economia nacional e segurança alimentar.

## **POTENTIAL IMPACT OF THIS RESEARCH**

This research offers significant benefits to society by improving the efficiency of genetic improvement programs for Gir and Girolando cattle breeds, contributing to the reduction of economic losses associated with lethal and sublethal recessive alleles. By identifying and mapping lethal haplotypes as well as segregating loci between breeds, this research provides tools to prevent matings that result in non-viable offspring, thereby promoting greater reproductive and productive efficiency. Additionally, the results can support actions within genetic improvement programs, such as the National Gir Dairy Breeding Program (PNMGL) and the Girolando Genetic Improvement Program (PMGG), directly impacting the economic and environmental sustainability of Brazilian dairy farming, a sector essential to the national economy and food security.

**CERTIFICADO DE APROVAÇÃO**

TÍTULO DA DISSERTAÇÃO: IDENTIFICAÇÃO DE HAPLÓTIPOS HOMOZIGOTOS PORTADORES DE POSSÍVEIS MUTAÇÕES LETAIS E SEMILETAIS RECESSIVAS EM POPULAÇÕES DE BOVINOS GIR E GIROLANDO


**AUTORA: BÁRBARA CAROLINA GOMES RIBEIRO**

**ORIENTADOR: FERNANDO SEBASTIAN BALDI REY**

**COORIENTADOR: MARCOS VINÍCIUS GUALBERTO BARBOSA DA SILVA**


**COORIENTADORA: ELISA PERIPOLLI**

Aprovada como parte das exigências para obtenção do Título de Mestra em Genética e Melhoramento Animal, pela Comissão Examinadora:


Documento assinado digitalmente  
 **MARCOS VINICIUS GUALBERTO BARBOSA DA SI**  
Data: 13/11/2024 10:38:58-0300  
Verifique em <https://validar.iti.gov.br>

Pesquisador Dr. MARCOS VINÍCIUS GUALBERTO BARBOSA DA SILVA (Participação Virtual)  
Centro Nacional de Pesquisa de Gado de Leite / EMBRAPA - Juiz de Fora/MG

Prof. Dr. RAFAEL ESPIGOLAN (Participação Virtual)  
Departamento de Zootecnia e Ciências Biológicas / Universidade Federal de Santa Maria - Palmeira das Missões/RS

Documento assinado digitalmente  
 **RAFAEL ESPIGOLAN**  
Data: 15/11/2024 12:27:02-0300  
Verifique em <https://validar.iti.gov.br>

Prof. Dr. FÁBIO LUIZ BURANELO TORAL (Participação Virtual)  
Departamento de Zootecnia / Universidade Federal de Minas Gerais (UFMG) - Belo Horizonte/MG

Documento assinado digitalmente  
 **FABIO LUIZ BURANELO TORAL**  
Data: 15/11/2024 01:49:15-0300  
Verifique em <https://validar.iti.gov.br>

Jaboticabal, 01 de novembro de 2024

## **DADOS CURRICULARES DO AUTOR**

**BÁRBARA CAROLINA GOMES RIBEIRO** - Nascida na cidade de Uberlândia, Minas Gerais, no dia 01 de janeiro de 1989, filha de Maria Aldair Gomes e José Cristóvão Gomes. Graduada em Engenharia Civil, em 2015, pelo Centro Universitário do Triângulo (Unitri), graduada em licenciatura plena em Matemática, em 2017, pela Universidade de Franca (Unifran) e graduanda do curso de Zootecnia pela Universidade Federal de Uberlândia (UFU). Na UFU, durante três semestres seguidos (2018 – 2019), foi monitora bolsista das disciplinas de Mecanização Agrícola e Melhoramento Genético Animal I e pelo período de um ano (2019 – 2020) foi integrante do Grupo de Estudos em Melhoramento Genético Animal (GEMEGA) responsável pela realização de provas de eficiência alimentar (EA) sob supervisão da Professora Doutora Carina Ubirajara. Foi membro parceiro da empresa júnior de consultoria veterinária (Conavet - UFU) prestando assistência técnica e consultoria para pequenas e médias propriedades leiteiras através da execução do Projeto Leite a Pasto sob supervisão do Professor Edmundo Benedetti (2019-2020). Realizou estágio extracurricular com bovinos leiteiros no Instituto de Zootecnia (IZ) na cidade de Nova Odessa – SP sob supervisão do pesquisador Dr. Luiz Carlos Roma Júnior; estágio extracurricular em confinamento da JHS Alimentos na cidade de Capinópolis-MG (2019-2020) e estágio curricular obrigatório de Assistência Técnica e Gerencial (ATeG - Senar) em parceria com B.F. Assessoria e Consultoria Agropecuária sob supervisão da Zootecnista Bianca Aparecida Freitas de Faria na cidade de Monte Alegre de Minas Gerais - MG. Ingressou em 2021 no programa de mestrado em Genética e Melhoramento Animal, na Faculdade de Ciências Agrárias e Veterinárias da Universidade Estadual Paulista - Unesp, Campus de Jaboticabal, sob a orientação do Professor Doutor Fernando Baldi e coorientação do Doutor Marcos Vinícius, com projeto de pesquisa em parceria com a Embrapa Gado de Leite e o apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

"Embora ninguém possa voltar atrás e fazer um novo começo, qualquer um pode começar agora e fazer um novo fim."

James R. Sherman

## **AGRADECIMENTOS**

À Deus, por me amparar em diversos momentos da caminhada e olhar por todos aqueles que amo.

Ao meu filho, Gabriel, por ressignificar nossas vidas e me mostrar o quão inexplicável e incondicional é o amor de mãe.

Ao meu marido, Bruno, pelo amor, admiração, dedicação e companheirismo. A vida com você é sempre uma soma e caminhar ao seu lado é partilhar propósitos e construir o futuro que sempre sonhamos.

Aos meus pais, Maria Aldair e José Cristóvão, por todo amor, confiança e carinho. Vocês são meu porto seguro e minha inspiração. À minha irmã, Ana e ao meu cunhado, Juliano, por todo companheirismo, amizade, motivação e presença.

Ao meu orientador, Dr. Fernando Baldi, pela oportunidade dada ao me selecionar como orientanda, pelos valiosos ensinamentos no decorrer da pós-graduação e por todo o entendimento e paciência que teve para comigo. Você é um pesquisador de excelência e é extraordinário dividir essa jornada com você.

Ao meu coorientador, Dr. Marcos Vinicius, pela disponibilidade de parceria nesta pesquisa, pelos valiosos ensinamentos e pelo profissionalismo.

Aos professores do curso de Graduação em Zootecnia da Universidade Federal de Uberlândia, em especial à professora Dra. Carina Ubirajara que me inspirou a especializar na área de melhoramento animal e é um exemplo de mulher, mãe e fortaleza. À você minha eterna admiração.

Aos professores do Programa de Pós-graduação de Genética e Melhoramento Animal pelos valiosos ensinamentos e à Universidade Estadual Paulista Júlio de Mesquita Filho – Faculdade de Ciências Agrárias e Veterinárias (FCAV/UNESP), pela oportunidade de cursar o programa de pós-graduação.

À Embrapa Gado de Leite, por conceder o banco de dados utilizado nesse estudo.

À Coordenação de Aperfeiçoamento Pessoal de Nível Superior (Capes) pela bolsa concedida durante o mestrado.

Gratidão é a palavra que resume tudo!

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

## SUMÁRIO

1.	INTRODUÇÃO .....	1
2.	REVISÃO DE LITERATURA.....	5
2.1	Variações genéticas e seleção genômica.....	5
2.2	Depressão por endogamia e alelos letais .....	6
2.3	Haplótipos.....	7
2.3.1	Identificação de haplótipos recessivos letais .....	8
2.4	Análise funcional: genes candidatos, análise de enriquecimento, genes ortólogos e fenótipos	11
3.	MATERIAL E MÉTODOS .....	13
3.1.	Imputação, filtragem, e faseamento de genótipos .....	13
3.2.	Testes para identificação de candidatos a haplótipos letais e semiletais .....	14
3.2.1	Teste populacional .....	15
3.2.2	Testes de acasalamento de portadores .....	15
3.2.3	Testes de semiletalidade .....	16
3.3	Seleção de um haplótipo como recessivo letal ou semiletal .....	17
3.4.	Identificação de genes candidatos, análise de enriquecimento funcional, genes ortólogos e fenótipos de letalidade.....	18
4.	RESULTADOS E DISCUSSÕES.....	18
4.1	Faseamento do genótipo e montagem dos haplótipos .....	18
4.2	Candidatos a haplótipos letais e semiletais identificados .....	19
4.3	Regiões candidatas e haplótipos selecionados que carregam supostos alelos recessivos letais ou semiletais .....	21
4.4	Genes candidatos, análise de enriquecimento funcional, genes ortólogos e fenótipos de letalidade .....	30
4.5	Discussões gerais.....	35
5.	CONCLUSÃO .....	39
6.	REFERÊNCIAS .....	41

## IDENTIFICAÇÃO DE HAPLÓTIPOS HOMOZIGOTOS PORTADORES DE POSSÍVEIS MUTAÇÕES LETAIS E SEMILETAIS RECESSIVAS EM POPULAÇÕES DE BOVINOS GIR E GIROLANDO

**RESUMO** – A produtividade e a lucratividade de um rebanho são impactadas principalmente pela eficiência reprodutiva. Todavia, a ocorrência de alelos recessivos deletérios pode resultar em grandes perdas econômicas em razão da geração de indivíduos homozigotos que resultarão em morte embrionária precoce ou logo após o nascimento. Objetivou-se, com este estudo, identificar haplótipos homozigotos portadores de mutações letais e semiletais nas raças Gir e Girolando, com base na frequência populacional, bem como identificar a segregação de loci recessivos letais entre as raças Gir e Girolando. Genótipos imputados de 64.251 animais das raças Gir e Girolando, foram submetidos ao controle de qualidade (CQ) e à inferência haplotípica através dos programas “pregsf90” e “AlphaPhase” V1.2, respectivamente. As estatísticas dos haplótipos foram calculadas por seis testes distintos: i) baseado na frequência de animais portadores ( $f_{Hh}$ ) e do número de animais genotipados (N), onde probabilidade de não observar nenhum animal homozigoto ( $P_{hh}$ ) foi calculada por  $P_{hh} = [(1 - f_{Hh}^2)/4]^N$ ; ii) teste de acasalamento de portadores para acasalamentos touro × matriz, onde probabilidade de não observar nenhum animal homozigoto ( $P_{hh}$ ) foi calculada por  $P_{hh} = 0,75^{n_s \times n_d}$ ; iii) teste de acasalamento de portadores para acasalamentos touro × avô materno, onde probabilidade de não observar nenhum animal homozigoto ( $P_{hh}$ ) foi calculada por  $P_{hh} = 0,875^{n_s \times n_{mgs}}$ ; iv) teste de semiletalidade para acasalamentos touro × matriz; v) teste de semiletalidade para acasalamentos touro × avô materno, e; vi) teste de semiletalidade populacional, onde para os três testes de semiletalidade foi calculada uma estatística de teste qui-quadrado unidirecional, com um grau de liberdade,  $\chi^2 = (O_{hh} - E_{hh})^2/E_{hh} + (O_{Hh + HH} - E_{Hh + HH})^2/E_{Hh + HH}$ . Genes candidatos foram identificados por meio do pacote “Bioconductor” versão 3.18 com o pacote “biomaRt” versão Ensembl Genes 103, sendo posteriormente confrontados com genes ortólogos, via pacote “orthogene”, e checados no banco de dados *Mouse Genome Informatics* para caracterização de possíveis efeitos letais ou semiletais dos haplótipos. Como resultado, 37 haplótipos (oito para raça Gir e vinte e nove para a raça Girolando) apresentaram redução de homozigosidade significativa. Dois haplótipos da raça Gir (GI1H2 e GI23H8) ocorreram em regiões semelhantes em estudos de identificação de haplótipos letais ou semiletais para outras raças taurinas e zebuínas, o que pode indicar uma conservação

de haplótipos do ancestral bovino. Duas regiões (GIGR3R1 e GIGR15R2) se sobrepueram parcialmente em ambas as raças. Treze genes candidatos presentes em 11 haplótipos (dois da raça Gir e nove na raça Girolando) se mostraram letais ou semiletais nas fases embrionária, pré-natal ou perinatal. Os genes *SYT1* (GI5H3), *PROC* (GR2H1) e *PIK3R1* (GR20H24) corroboraram para a provável letalidade e semiletalidade dos respectivos haplótipos, bem como os genes *SPTA1* (região GIGR3R1) e *SOX6* (região GIGR3R2) corroboraram para uma possível segregação alélica semiletal. Os resultados têm aplicação prática nos programas de melhoramento genético das raças Gir (Programa Nacional de Melhoramento do Gir Leiteiro – PNMGL) e Girolando (Programa de Melhoramento Genético do Girolando – PMGG), permitindo identificar animais portadores desses haplótipos e incluir essas informações nos sumários de touros. Isso possibilita aos criadores evitarem acasalamentos que resultem em progênes inviáveis, contribuindo para maior eficiência reprodutiva na bovinocultura leiteira.

**Palavras – chaves:** Alelos recessivos, Endogamia, Inferência de haplótipos, Fertilidade, Indivíduos homocigotos.

## IDENTIFICATION OF HOMOZYGOUS HAPLOTYPES CARRYING POSSIBLE RECESSIVE LETHAL AND SEMILETHAL MUTATIONS IN GIR AND GIROLANDO CATTLE POPULATIONS

**ABSTRACT** – Productivity and profitability in a herd are primarily influenced by reproductive efficiency. However, the occurrence of deleterious recessive alleles can lead to significant economic losses due to the generation of homozygous individuals that result in early embryonic death or death shortly after birth. This study aimed to identify homozygous haplotypes carrying lethal and sublethal mutations in Gir and Girolando cattle breeds based on population frequency, as well as to detect the segregation of lethal recessive loci between the Gir and Girolando breeds. Imputed genotypes from 64,251 animals of the Gir and Girolando breeds were subjected to quality control (QC) and haplotype inference using the “pregsf90” and “AlphaPhase” V1.2 programs, respectively. Haplotype statistics were calculated using six distinct tests: i) based on the frequency of carrier animals ( $f_{Hh}$ ) and the number of genotyped animals (N), where the probability of not observing any homozygous animal (Phh) was calculated as  $Phh = [(1 - f_{Hh}^2)/4]^N$ ; ii) mating test for carrier sire × dam matings, where the probability of not observing any homozygous animal (Phh) was calculated by  $Phh = 0,75^{n_s \times n_d}$ ; iii) mating test for carrier sire × maternal grandsire matings, where the probability of not observing any homozygous animal (Phh) was calculated by  $Phh = 0,875^{n_s \times n_{mgs}}$ ; iv) semilethality test for sire × dam matings; v) semilethality test for sire × maternal grandsire matings, and; vi) population semilethality test, where for the three semilethality tests a unidirectional chi-square test statistic was calculated, with one degree of freedom,  $\chi^2 = (O_{hh} - E_{hh})^2/E_{hh} + (O_{HH} + HH - E_{HH} + HH)^2/E_{HH} + HH$ . Candidate genes were identified using the “Bioconductor” package version 3.18 with the “biomaRt” package (Ensembl Genes 103 version), subsequently compared with orthologous genes via the “orthogene” package, and cross-checked with the Mouse Genome Informatics database to characterize potential lethal or sublethal effects of the haplotypes. As a result, 37 haplotypes (eight in the Gir breed and twenty-nine in the Girolando breed) exhibited significant reductions in homozygosity. Two haplotypes in the Gir breed (GI1H2 and GI23H8) were found in regions similar to those identified in studies of lethal or sublethal haplotypes in other taurine and zebu breeds, suggesting conservation of ancestral bovine haplotypes. Two regions (GIGR3R1 and GIGR15R2) partially overlapped in both breeds. Thirteen candidate genes present in 11 haplotypes (two in the Gir breed and nine in the Girolando

breed) were found to be lethal or sublethal during embryonic, prenatal, or perinatal stages. The genes *SYT1* (GI5H3), *PROC* (GR2H1), and *PIK3R1* (GR20H24) supported the probable lethality and sublethality of the respective haplotypes, while the genes *SPTA1* (region GIGR3R1) and *SOX6* (region GIGR3R2) corroborated a possible sublethal allelic segregation. The results have practical applications in genetic improvement programs for the Gir (National Gir Dairy Breeding Program – PNMGL) and Girolando (Girolando Genetic Improvement Program – PMGG) breeds, enabling the identification of animals carrying these haplotypes and incorporating this information into bull summaries. This allows breeders to avoid matings that would result in non-viable progeny, contributing to improved reproductive efficiency in dairy cattle farming.

**Keywords:** Recessive alleles, Endogamy, Haplotype phasing, Fertility, Homozygous individuals.

## LISTA DE TABELAS

Tabela 1: Quantidade de haplótipos com ausência significativa ou nível reduzido de homozigidade por teste e raça. ....	19
Tabela 2: Haplótipos identificados carregando supostos alelos recessivos letais ou semiletais para as raças Gir e Girolando, e seus respectivos testes significativos. ....	22
Tabela 3: Estatísticas para os supostos haplótipos letais e semiletais para as raças Gir e Girolando.....	25
Tabela 4: Regiões identificadas carregando supostos haplótipos letais ou semiletais em possível segregação alélica da raça Gir para a raça Girolando. ....	28
Tabela 5: Estatísticas dos supostos haplótipos letais ou semiletais em segregação entre as raças Gir e Girolando. ....	29
Tabela 6: Descrição do termo MeSH (categoria D) e parâmetro gerado pelo teste hipergeométrico associado aos haplótipos supostamente letais ou semiletais para raça Gir. ....	31
Tabela 7: Fenótipo letal para genes presentes nos supostos haplótipos letais ou semiletais para a raça Gir.....	32
Tabela 8: Fenótipo letal para genes presentes nos supostos haplótipos letais ou semiletais para a raça Girolando.....	33
Tabela 9: Fenótipo letal para genes presentes nas regiões sobrepostas abrigando haplótipos de segregação alélica letal ou semiletal para a raça Girolando. ....	35

## LISTA DE TABELAS SUPLEMENTARES

Tabela Suplementar S1: Lista de regiões candidatas apresentando pelo menos cinco haplótipos com deficiência homozigótica para a raça Gir. ....	53
Tabela Suplementar S2: Lista de regiões candidatas apresentando haplótipos com deficiência homozigótica para a raça Girolando.....	55
Tabela Suplementar S3: Lista de regiões que possivelmente abrigam haplótipos supostamente letais ou semiletais em segregação da raça Gir para a raça Girolando.	56
Tabela Suplementar S4: Genes candidatos encontrados via recurso de anotação e análise de enriquecimento MeSH para haplótipos da raça Gir.....	57
Tabela Suplementar S5: Descrição do termo MeSH (categoria D) e parâmetro gerado pelo teste hipergeométrico associado aos haplótipos supostamente letais ou semiletais para raça Girolando.....	58
Tabela Suplementar S6: Genes candidatos encontrados via recurso de anotação e análise de enriquecimento MeSH para haplótipos da raça Girolando.....	62

## 1. INTRODUÇÃO

Segundo dados da Organização das Nações Unidas para Alimentação e Agricultura (FAO, 2023), a produção de leite brasileira ocupa o quinto lugar do ranking mundial, com potencial de expansão de produtividade para o mercado global de leite e derivados. Em nível nacional, a cadeia leiteira é uma das principais atividades econômicas e de geração de emprego e renda, sendo observada em quase todos os municípios brasileiros (Rocha et al., 2020). Os sistemas de produção de leite a pasto no Brasil requerem animais mais adaptados às condições climáticas do país, sendo observado nestes sistemas o cruzamento de animais das espécies *Bos primigenius taurus* e *Bos primigenius indicus* (Canaza-cayo et al., 2017). A raça Holandesa (*Bos primigenius taurus*) é considerada a mais expressiva raça leiteira do mundo em termos de volume de produção de leite (Albertino, 2021), ao passo que a raça Gir (*Bos primigenius indicus*), reconhecida como uma das raças zebuínas de maior produtividade leiteira em clima tropical e subtropical, se destaca em virtude de sua rusticidade, alta produtividade e maior longevidade no rebanho (Ospina, 2020). Nestes aspectos, com o intuito de alinhar a alta capacidade de produção de leite dos bovinos da raça Holandesa e a rusticidade da raça Gir para gerar progênes mais adaptadas e produtivas em climas tropicais e subtropicais, o cruzamento dessas duas raças vem sendo praticado no Brasil desde a década de 1940 (Silva et al., 2023).

Animais da raça Girolando, ou animais que possuam alguma fração da composição racial entre o Gir e o Holandês, são responsáveis por cerca de 80% da produção total de leite no Brasil. Essa raça é a que mais cresce na produção de sêmen no país, com mais de 920.848 doses produzidas em 2021, apresentando ao longo de 21 anos (2000-2021) um aumento de 85% em sua produção de leite em até 305 dias de lactação (3.683 kg/vaca para 6.821 kg/vaca) (Silva et al, 2023). Embora essas informações sejam consideradas positivas para o setor de produção leiteira no Brasil, ainda existem obstáculos a serem superados, como aqueles relacionados às características de importância econômica e seus impactos nos programas de melhoramento genético (Canaza-cayo et al., 2017).

O sucesso dos programas de melhoramento genético animal depende do objetivo e critérios de seleção. Enquanto o objetivo de seleção engloba as características que devem ser consideradas para gerar um ganho econômico para o

criador, sendo frequentemente expresso por meio de um índice econômico, os critérios de seleção são as características de importância econômica que serão efetivamente avaliadas para realizar a escolha dos animais (Zadra, 2012). As características de produção de leite e a taxa de fertilidade representam dois dos principais critérios de seleção em bovinos leiteiros (Albertino, 2021). Contudo, diversos estudos indicaram que a seleção genética para aumento da produção de leite pode reduzir a fertilidade, pois as características de produção e reprodução frequentemente apresentam correlações genéticas desfavoráveis (Ojango e Pollot, 2001; Pryce et al., 2002; Makgahlela et al., 2007; Sewalem et al., 2010).

Analisando outro aspecto, a seleção de animais geneticamente superiores, associada à utilização de biotecnologias reprodutivas como a inseminação artificial (IA) e a produção in vitro de embriões (PIVE), além de contribuir para o progresso genético, favorecem inevitavelmente o aumento da endogamia e a redução da diversidade genética (Canaza-cayo et al., 2014). A endogamia é definida como o acasalamento entre indivíduos geneticamente próximos ou aparentados, resultando, conseqüentemente, no aumento da homozigose, podendo apresentar implicações ao rebanho como aparecimento de mutações letais e subletais, bem como mutações que afetam a fertilidade (Adams et al., 2016). Neste sentido, embora o acasalamento entre indivíduos aparentados favoreça uma maior difusão de material genético superior, essa técnica também pode aumentar a frequência de alelos deletérios na população e a manifestação de fenótipos letais nas progênes (Wu et al., 2020).

No Gir Leiteiro, acasalamentos entre animais importados para a criação das primeiras linhagens resultaram em aumento da endogamia na população (Ospina, 2020), sendo constatado o aumento significativo da taxa de endogamia por geração nos estudos de Santana Júnior et al. (2016). O coeficiente de endogamia, que representa a probabilidade de dois alelos de um indivíduo serem idênticos por descendência (Wright, 1923), é uma métrica amplamente utilizada para avaliar o impacto dessa prática. Em bovinos da raça Gir, Reis Filho et al. (2010) encontraram valores de coeficiente de endogamia médio de 2,82%, enquanto Canaza-Cayo et al. (2014) encontraram coeficiente de endogamia médio de 0,11% para raça Girolando.

Tradicionalmente, o coeficiente de endogamia é estimado com base em informações de pedigree. No entanto, os avanços na genômica, por meio de painéis

de SNP de alta densidade, permitem estimar a endogamia ( $F$ ) de forma mais precisa, utilizando genótipos homozigotos (RHO, do inglês '*Runs of homozygosity*') com base no comprimento das regiões contínuas do genoma em homozigoses (ROH) (Ferencakovic et al., 2011). Em estudos sobre população de Gir Leiteiro, Peripolli et al. (2018) observaram um valor médio de endogamia genômica ( $F_{ROH}$ ) igual a 2,3% para uma população ancestral de três gerações, enquanto Ospina (2020) relatou uma  $F_{ROH}$  de 1,6%.

Esses avanços genômicos destacam o potencial da genômica como uma ferramenta indispensável em programas de melhoramento genético. Além de permitir a identificação de níveis mais precisos de endogamia, a genômica possibilita o conhecimento detalhado de genes e haplótipos (blocos de alelos de locos ligados, geralmente herdados como uma unidade). Esses elementos podem ser utilizados como marcadores de características produtivas, auxiliando na definição de estratégias de seleção genética para gado leiteiro (Gutierrez-reinoso et al., 2021).

Adicionalmente, a ampla disponibilidade de dados genômicos tem permitido identificar haplótipos relacionados a alelos recessivos letais. No estudo de VanRaden et al. (2011a), observou-se que certos haplótipos, comuns na população, não apresentavam homozigotos em animais vivos, caracterizando a presença de alelos letais. Essa metodologia se destacou por não depender de dados fenotípicos, como praticado nas metodologias de Lander e Botstein (1987), Remington e O'Malley (2000), Charlier et al., (2008), Khatib et al., (2009) e Huqun et al., (2010).

A metodologia de VanRaden et al. (2011a) tem se mostrado muito bem sucedida para criadores da raça Holandesa, visto que o teste de haplótipos pode ser usado para evitar acasalamentos entre portadores de alelos relacionados a defeitos genéticos e reduzir as frequências futuras dos alelos letais (Wu et al., 2020). Neste sentido, a partir de agosto de 2011, algumas associações de raças de bovinos leiteiros dos Estados Unidos (Holandês, Jersey e Pardo-Suíço) relataram o status de portador para todos os animais genotipados com o chip BovineSNP50 (Schmidit, 2020). Algo semelhante é observado nos Sumários de Touros Girolando, que apresentam o status de portador tanto para marcadores moleculares de interesse (Kappa-caseína, Beta-caseína e Beta-lactoglobulina), quanto para marcadores moleculares de doenças hereditárias: complexo de malformação vertebral – CVM, deficiência de adesão de

leucócitos bovinos – BLAD, deficiência da sintase de monofosfato de uridina – DUMPS, síndrome *Brachyspina* – BS e deficiência do Fator XI – FXID; para touros participantes do Teste de Progênie (Silva et al., 2023).

Embora tenha sido anunciada a inclusão das análises para doenças hereditárias nos Sumários de Touros Gir Leiteiro a partir de 2013, a última edição (Panetto et al., 2023) não apresenta o status de portador para CVM, BLAD e DUMPS. Atualmente essas doenças são consideradas ausentes no rebanho nacional da raça Gir, sendo que qualquer animal identificado como portador de algum alelo para estas doenças deve ser descartado (Panetto et al., 2023). Isto explicaria a falta de tais informações na última edição do Sumário de Touros Gir Leiteiro. Entretanto, ressalta-se que o status de portador para marcadores moleculares de interesse: Kappa-caseína e Beta-lactoglobulina (a partir de 2006), e Beta-caseína (a partir de 2016), são informados nos Sumário de Touros Gir Leiteiro para touros participantes do Teste de Progênie (Panetto et al., 2023). Entretanto, outras letalidades ainda não conhecidas podem existir

Considerando que o Brasil possui o segundo maior rebanho de vacas ordenhadas do mundo (FAO, 2019) e que 80% da produção nacional, que confere ao país o 5º lugar no ranking mundial, é advinda de gado Girolando (Silva et al., 2023), na qual a raça Gir possui papel fundamental em sua formação, torna-se importante obter informações sobre a ocorrência de haplótipos recessivos na população para auxílio na seleção genética e acasalamento dos animais. Nestas perspectivas, a presente pesquisa justifica-se por apresentar contribuições para estudos genômicos, especificamente, aos estudos de identificação de haplótipos recessivos letais nas raças bovinas Gir e Girolando, visto que se pode assumir a possível ocorrência de tais haplótipos nestas populações baseado nos resultados de pesquisas em outras raças bovinas de corte e leite.

Assim, o objetivo desta pesquisa foi identificar e mapear haplótipos letais e semiletais específicos das raças Gir e Girolando, bem como identificar a segregação para a raça Girolando de loci recessivos letais encontrados, neste estudo, na raça Gir. A partir do objetivo geral, listam-se como objetivos específicos desta pesquisa: i) Realizar a imputação, filtragem e faseamento dos genótipos; ii) Identificar candidatos a haplótipos letais e haplótipos semiletais na raça Gir; iii) Identificar candidatos a

haplótipos letais e haplótipos semiletais na raça Girolando; iv) Verificar a ocorrência de segregação de haplótipos letais para a raça Girolando, estes já identificados na raça Holandesa e os apresentados na raça Gir, e; v) Identificar genes candidatos e realizar a análise funcional para os haplótipos encontrados nas raças Gir e Girolando.

## **2. REVISÃO DE LITERATURA**

### **2.1 Variações genéticas e seleção genômica**

O fenótipo de um indivíduo é definido pelo seu código genético completo, isto é, a sequência do seu genoma, sob influência do ambiente. Variações genéticas nas sequências do genoma contribuem para variações fenotípicas observadas em uma população (Lehner, 2013). Essas variantes genéticas podem ser substituições de um único nucleotídeo que ocorre em uma posição específica no genoma (SNPs), ou deleções ou inserções curtas (indels) (Derks, 2020).

Neste sentido, a seleção de animais com melhor desempenho para serem pais da próxima geração, baseado nas variações genéticas, tornou-se um caminho pertinente (Georges et al., 2019), mas considerado oneroso e demorado até pouco tempo atrás. Entretanto, o avanço de ferramentas de sequenciamento do genoma possibilitou o uso dessas variações genéticas para a seleção de animais.

Até recentemente, os genomas de algumas espécies animais eram de qualidade inferior se comparado com o genoma humano ou do camundongo de laboratório. Com o advento da tecnologia de sequenciamento de leitura longa, foram aprimorados genomas de referência para diversas espécies, melhorando significativamente a descoberta de variantes genéticas e análises genéticas, como por exemplo o genoma de galinha (Warren et al., 2017), o genoma suíno (Warr et al., 2020) e o genoma de referência bovino *Bos taurus* ARS-UCD 1.2 (Rosen et al., 2020).

A seleção genômica, que utiliza chips de genotipagem, permitiu um grande avanço no melhoramento animal. Os chips de SNPs permitem capturar as variações genéticas dentro de uma raça, trazendo qualidade em informações genômicas, usadas para selecionar os melhores animais para reprodução (Dekkers, 2012). Entretanto, a seleção genômica não é muito eficiente na eliminação de variações

deletérias raras porque os alelos de baixa frequência não são bem capturados pelos SNPs nos chips de seleção (Zhang et al. 2019).

## **2.2 Depressão por endogamia e alelos letais**

Segundo Bosse et al. (2018), a domesticação de espécies tem impacto importante em termos de tamanho efetivo da população e pressão de seleção, o que poderia diminuir o “pool” genético. Ainda segundo esses autores, populações reduzidas são mais propensas a aumentar os efeitos do sistema de acasalamento endogâmico, levando a progênie endogâmica ao declínio de aptidão, chamado de depressão por endogamia. A depressão por endogamia é o resultado do acúmulo de alelos deletérios que aumentam em frequência (Charlesworth e Willis, 2009), ou seja, o acúmulo de mutações recessivas prejudiciais no genoma, pois a endogamia aumenta a probabilidade dessas mutações se tornarem homozigotas (Agrawal e Whitlock, 2012).

Mutações embrionárias letais são a manifestação mais precoce e talvez a mais grave da depressão por endogamia (Stoffel et al., 2023). É provável que sejam relativamente comuns, uma vez que as mutações de perda de função são letais em cerca de um terço dos genes dos mamíferos, e a maioria destas são provavelmente letais no pré-natal, e não no pós-natal (Dickinson et al., 2016; Georges et al., 2019).

Os efeitos letais recessivos são expostos apenas em cruzamentos de animais portadores. Segundo Berry e Spangler (2023), algumas associações de criadores recusam-se a registrar indivíduos portadores de alelos para variantes específicas. Na raça Senepol, o gene “SLICK”, relacionado a capacidade animal de tolerar altas temperaturas do ambiente, e o gene “Dupla Musculatura”, responsável pelo desenvolvimento anormal da massa muscular em bovinos, são exemplos para negativas de registro (Menezes et al., 2022). Reprodutores portadores dos genes homozigotos indesejáveis para as respectivas características (ss - homozigoto com ausência do gene Slick e DMDM - homozigoto para Dupla Musculatura) não podem ser inscritos para comercialização de sêmen, em centros de coleta e processamento de sêmen (CCPS) registrados no Ministério da Agricultura, Pecuária e Abastecimento (MAPA).

O conhecimento do status de portador de animais individuais também é informativo ao decidir sobre acasalamentos para produzir a próxima geração. Neste sentido, a indicação de status portador de alelos para doenças hereditárias é encontrada no Teste de Progênie do Sumário Brasileiro de Touros Gir Leiteiro e no Teste de Progênie do Sumário de Touros do Programa de Melhoramento Genético da Raça Girolando, apresentando os marcadores para doenças hereditárias como Complexo de Má-formação Vertebral (CVM), Deficiência de Adesão Leucocitária Bovina (BLAD), Deficiência da Uridina Monofosfato Sintetase (DUMPS), *Brachyspina* (BS) e Deficiência do Fator XI (FXID) (Panetto et al., 2023; Silva et al., 2023).

Com o surgimento de novas tecnologias de sequenciamento, o impacto da depressão endogâmica e de alelos deletérios, pode ser avaliado. Atualmente, a identificação de alelos deletérios a partir de dados de Seleção Genômica Ampla (GWS) está voltada para variações de proteínas que afetam a sequência codificadora do gene. Segundo Derks (2020), softwares como o *Protein Variation Effect Analyzer - PROVEAN* (Choi e Chan, 2015) e algoritmos como o *Sorting Intolerant From Tolerant - SIFT* (Kumar et al., 2009) são frequentemente usados para avaliar variantes missense deletérias, causadas por uma única alteração de nucleotídeo que resulta em um códon que codifica um aminoácido diferente na proteína.

### 2.3 Haplótipos

O genoma de uma espécie é altamente similar entre diferentes indivíduos de uma mesma população, sendo conservadas regiões específicas na sequência de DNA (ácido desoxirribonucleico) através das gerações. O conjunto de alelos em locos adjacentes, que fazem parte de um mesmo cromossomo e são transmitidos em blocos através das gerações, é denominado de haplótipo (Nih, 2024; Derks, 2020).

Segundo Cole (2016), nos Estados Unidos (EUA), aproximadamente 11 milhões de dólares em perdas econômicas poderiam ser evitadas se acasalamentos realizados prevenissem a formação de embriões acometidos por haplótipos indesejáveis, evitando a redução da fertilidade e morte perinatal de bezerras. A identificação de haplótipos através de dados genômicos é uma ferramenta estratégica para a manutenção genética de um rebanho, já que alguns haplótipos trazem

consequências negativas, incluindo redução da lucratividade por sua propagação no rebanho (STgenetics®, 2024).

Em relação a divisão do genoma em blocos de haplótipos, há dois métodos para inferência dos haplótipos e que são empregados em diversos softwares (Araújo, 2022): i) o método que utiliza comprimento fixo (Villumsen et al., 2009), onde o genoma é dividido de acordo com o número de SNPs no bloco ou pelo comprimento em megabases, e; ii) o método que utiliza comprimento variável (Rinaldo et al., 2005), que considera o desequilíbrio de ligação (LD - associação não aleatória entre dois locos que pode ser afetado pela diversidade genética e estrutura populacional dos rebanhos) para montagem de haplótipos de tamanhos variados.

Os programas “FImpute” (Sargolzaei et al., 2014), “Beagle” (Browning e Browning, 2009) e “findhap” (VanRaden et al., 2011c) são exemplos de ferramentas que utilizam o método de comprimento variável através das janelas deslizantes (*sliding window*) enquanto os programas “AlphaPhase” (Hickey et al., 2011) e “SHAPEIT” (Delaneau et al., 2011) utilizam o método de comprimento fixo. Cada método de reconstrução de haplótipos apresenta vantagens e desvantagens. O método fixo tem como vantagem a facilidade de se obter os haplótipos (Hess et al., 2017), enquanto o método variável contabiliza a co-segregação dos alelos através da identificação de *hotspots* de recombinação, além de considerar o LD. As desvantagens, no método fixo, seriam não consideram o LD ou pontos de recombinação no interior dos haplótipos; e demandar maior tempo de construção e serem específicos de uma determinada população no método variável (Araújo, 2022; Cuyabano et al., 2014).

### **2.3.1 Identificação de haplótipos recessivos letais**

Existem diferentes métodos para a detecção de alelos letais durante a gestação de um embrião, sendo uma delas o uso de haplótipos recessivos para identificação de letalidades e informação de status portador de reprodutores e matrizes em populações em acasalamento (Stoffel et al., 2023).

Estudos de vários haplótipos recessivos em raças leiteiras relacionados, por exemplo, a doenças hereditárias, reprodução e fertilidade, são relatados na literatura.

Em bovinos da raça Holandesa, os haplótipos recessivos HH1, HH2, HH3, HH4 e HH5, que afetam a fertilidade e a ocorrência de natimortos, foram descritos por Adams et al. (2016), McClure et al. (2014), VanRaden et al. (2011b), Fritz et al. (2013) e Cooper et al. (2013). Fritz et al. (2018) relataram a descoberta do haplótipo HH6 relacionado à morte embrionária na raça Holandesa, na França. Os fatores causais subjacentes em alguns desses haplótipos recessivos são relatados em estudos de diversos autores: haplótipo HH1 apresenta uma mutação sem sentido (substituição de nucleotídeos que resulta em um códon de parada e, conseqüentemente, no término prematuro da proteína, resultando na produção de uma proteína encurtada que, provavelmente, não é funcional) no gene APAF1 (Adams et al., 2016); haplótipo HH3 apresenta uma mutação não-sinônima (substituição de nucleotídeos que leva a uma substituição de aminoácidos, podendo ou não resultar em uma variante patogênica dependendo do efeito da substituição de aminoácidos na função e estrutura da proteína) no gene SMC2 (McClure et al., 2014); haplótipo HH4 apresenta uma mutação não-sinônima no gene GART (Fritz et al., 2013); haplótipo HH5 apresenta uma deleção de 138 kbp no gene TFB1M (Schütz et al., 2016); e haplótipo HH6 apresenta uma mutação de códon iniciador no gene SDE2 (Fritz et al., 2018).

Em relação aos haplótipos recessivos ligados a doenças hereditárias, o *Council on Dairy Cattle Breeding* (CDCB) relatou em bovinos da raça Holandesa, mutações para: complexo de malformação vertebral – CVM (haplótipo HHC), deficiência de adesão de leucócitos bovinos – BLAD (haplótipo HHB), deficiência da sintase de monofosfato de uridina – DUMPS (haplótipo HDD), *Brachyspina* – BY (haplótipo HH0), sindactilia (*mule foot*) – MF (haplótipo HHM), e deficiência de colesterol – CD (haplótipo HCD).

Tratando-se de outras raças, os estudos de Cooper et al. (2014) e Null et al. (2017) identificaram, em bovinos Ayrshire, um haplótipo que afeta a taxa de concepção (AH1) e um haplótipo que afeta a fertilidade (AH2), respectivamente. Em relação a fatores causais subjacentes a haplótipos recessivos em outras raças leiteiras, são também descritos: no haplótipo JH1 de bovinos Jersey observa-se uma mutação nonsense no gene CWC15 (Sonstegard et al., 2013); em bovinos da raça Pardo-Suíço, no haplótipo recessivo BH2, observa-se uma mutação de sentido trocado no gene TUBD1 (Schwarzenbacher et al., 2016); e em gado Normando, o

haplótipo recessivo NH2 apresenta uma mutação de sentido trocado no gene CAD (Mesbah Uddin et al., 2019). Haplótipos recessivos ligados a doenças hereditárias, segundo o *Council on Dairy Cattle Breeding* (CDCB), são relatados em bovinos Pardo-Suíço: desmielinização espinhal – SMD (haplótipo BHD), atrofia muscular espinhal – SMA (haplótipo BHM) e Síndrome de Weaver – WS (haplótipo BHW).

VanRaden et al. (2011a) propuseram uma metodologia simples para identificação de haplótipos letais. Segundo os autores, o número esperado de indivíduos homocigotos para qualquer haplótipo específico podem ser calculados de duas maneiras: i) método “simples” que utiliza o número de indivíduos genotipados dividido por quatro e multiplicado pelo quadrado da frequência de portadores; ii) método padrão de acasalamento real, que verifica o número de indivíduos genotipados na população descendentes de tais acasalamentos, e que é calculado como o número de acasalamentos de touro portador x avô materno portador (MGS) dividido por quatro. Neste método, as frequências alélicas para avós maternas são assumidas como iguais às dos MGS. Ressalta-se que as expectativas do método simples podem ser superestimadas se a endogamia for evitada ou subestimadas se as frequências alélicas mudarem ao longo do tempo, enquanto as expectativas do método padrão de acasalamento real pode ser subestimadas se alguns touros portadores e MGS portadores não forem genotipados (VanRaden et al., 2011a).

As probabilidades de observar nenhum homocigoto quando “n” são esperados podem ser obtidas por duas fórmulas análogas às usadas para obter as expectativas: i) para o método simples, a probabilidade é dada por 1 menos a frequência portadora ao quadrado dividido por quatro, elevada à potência do número de animais genotipados; ii) para o método padrão de acasalamento real, a probabilidade é igual a 0,75 elevado à potência do número observado de acasalamentos touro portador x MGS portador (VanRaden et al., 2011a).

Jenko et al. (2019) propuseram uma diferenciação para o método padrão de acasalamento real de VanRaden et al. (2011a), ao apresentarem o número de descendentes tanto para os cruzamentos de touro portador x MGS portador, quanto para os cruzamentos touro portador x matriz portadora. Nesse estudo, o número de descendentes genotipados dos cruzamentos de touro portador x MGS portador é calculado como o número de acasalamentos de touro portador x MGS portador

dividido por oito, enquanto o número de descendentes genotipados dos cruzamentos de touro portador x matriz portadora é calculado como o número de acasalamentos de touro portador x matriz portadora dividido por quatro. Para as probabilidades de observar nenhum homozigoto quando “n” são esperados, que seguem um processo de Bernoulli, são calculadas como: 0,875 elevado à potência do número observado de acasalamentos touro portador × MGS portador e 0,75 elevado à potência do número observado de acasalamentos touro portador × matriz portadora, respectivamente (Jenko et al., 2019).

A metodologia de VanRaden et al. (2011a) é empregada em estudos de haplótipos recessivos em bovinos de corte (Pausch et al., 2015; Hoff et al., 2017; Jenko et al., 2019; Schmidit et al., 2023), em suínos (Häggman e Uimari, 2017; Howard et al., 2017; Derks et al., 2017), em ovelhas leiteiras (Ben Braiek, et al., 2021) e, recentemente, em ovelhas selvagens (Stoffel et al., 2023). Assim, percebe-se que os testes de haplótipos recessivos são usados rotineiramente em animais genotipados, de espécies diversificadas, para descobrir novas variantes deletérias e identificar animais portadores.

#### **2.4 Análise funcional: genes candidatos, análise de enriquecimento, genes ortólogos e fenótipos**

Segundo Sollero e Grynberg (2020), dados biológicos advindos do conhecimento genômico são relativamente complexos em função da quantidade excessiva de informações, visto que são uma lista de marcadores, haplótipos ou genes significativos, associados a uma análise estatística. As pesquisas sobre classificação funcional são relevantes quando mostram importância em relação ao assunto em questão, os termos utilizados estão alinhados com o contexto biológico e há a possibilidade de identificar as vias metabólicas e proteicas que têm um papel conhecido na expressão ou variação de uma característica analisada (Sollero e Grynberg; 2020).

O projeto Bioconductor oferece ferramentas para a análise e compreensão desses dados partindo de uma extensa lista de recursos para mapeamento de marcadores SNPs, incluindo vias metabólicas, Ontologia genética (Gene Ontology -

GO) e anotações, que podem ser realizadas pelo pacote biomaRT, através do banco de dados Ensembl e via recursos de anotações como o MeSH (Medical Subject Headings) (Sollero e Grynberg; 2020).

MeSH é uma coletânea de mais de 25.000 registros clínicos e biológicos para 70 espécies, incluindo categorias não tradicionalmente representadas em GO. Os mais de 4.000 termos MeSH são definidos em artigos indexados pela *National Library of Medicine* no PubMed e podem ser mapeados diretamente em genes para desenvolvimento de anotação genética (Sollero e Grynberg; 2020). A aplicação desta ferramenta tem sido utilizada em diversos estudos com animais domésticos e espécies vegetais, como por exemplo, nos estudos de Oliveira Júnior et al., 2019; Morota et al., 2016; Morota et al., 2015, entre outros. De acordo com Sollero e Grynberg (2020), entre as categorias utilizadas no MeSH no contexto animal, cinco delas são mais importantes: A (Anatomia), B (Organismos), C (Doença), D (Drogas e Químicos) e G (Fenômeno e Processo).

A análise de enriquecimento funcional via MeSH, é realizada através das seguintes etapas apresentadas por Sollero e Grynberg (2020): i) identificação da lista de genes de interesse (ENTREZID), ii) acesso à anotação do genoma e extração de lista de marcadores mais informativos, que será uma lista “background” de genes, e iii) teste com significância estatística para identificar genes, vias e, ou, funções metabólicas que estão enriquecidos.

Conforme Morota et al (2015), o grau de enriquecimento MeSH é calculado como uma probabilidade que indica quais termos específicos foram detectados com mais frequência do que o esperado ao acaso e o teste hipergeométrico é usado para calcular essas probabilidades. Termos identificados como importantes são extraídos para formar redes, compreender vias genéticas ou metabólicas e/ou formular novas hipóteses, sendo que o p-valor representa o nível de significância de cada termo representado pelos genes, enquanto o p-ajustado ou q-valor é a correção do p-valor para testes de múltiplas hipóteses (Morota et al., 2016).

Outros mecanismos, como a revisão de literatura por exemplo, são necessários para definir quais termos apontados pelo MeSH podem explicar mecanismos genéticos ou sugestões de alguns possíveis genes candidatos relacionados às características estudadas (Sollero e Grynberg; 2020). Nesta perspectiva, o pacote

“orthogene” é uma ferramenta de fácil mapeamento de genes ortólogos atualizados de mais de 700 organismos (Schilder e Skene, 2022), permitindo a consulta de informações funcionais entre espécies distintas. Em complemento, o banco de dados do genoma do camundongo (*Mouse Genome Database - MGD*), principal fonte de dados para camundongos de laboratório oferecendo informações sobre genes, marcadores genéticos e características genômicas, apresenta-se como uma ferramenta para consulta de fenótipos, informações funcionais, entre outras possibilidades (Blake et al., 2017).

### 3. MATERIAL E MÉTODOS

#### 3.1. Imputação, filtragem, e faseamento de genótipos

Os dados genotípicos utilizados nesta pesquisa foram inicialmente compostos por 35.550 animais da raça Gir e 28.696 animais da raça Girolando, que foram genotipados com painéis do tipo SNP da Illumina BovineHD BeadChip (San Diego, CA, EUA) e GeneSeek® Genomic Profiler *indicus* (GGP *indicus* 50K), respectivamente. Segundo Hickey et al. (2012), como os marcadores nos painéis de média e alta densidade tendem a apresentar forte desequilíbrio de ligação, reduzindo assim os erros de imputação, a densidade de marcadores se torna importante fator com influência direta na acurácia.

Os animais genotipados com painel de menor densidade foram imputados diretamente ao painel Illumina BovineHD BeadChip através do software “FImpute” versão 2.2 (Sargolzaei et al., 2014). A etapa de imputação foi realizada previamente pela Embrapa Gado de Leite, sendo, portanto, repassados para continuidade desta pesquisa os genótipos de 35.552 animais da raça Gir e 28.699 animais da raça Girolando, com 420.722 e 396.862 marcadores SNP autossômicos das raças Gir e Girolando, respectivamente.

O controle de qualidade (CQ) dos dados genômicos foi realizado através do programa “pregsf90” (Aguilar et al., 2010). Dentre os critérios de CQ para os SNPs e seus limites de exclusão, foi considerado a proporção de SNPs identificados (*call rate*)  $\geq 0,90$ , a frequência mínima de alelos (MAF)  $\geq 0,05$ , teste estatístico do Equilíbrio de

Hardy-Weinberg (HWE) e exclusão de SNPs monomórficos. Dentre os critérios de CQ para as amostras e seus limites de exclusão, foi considerado call rate  $\geq 0,90$  e removidos progênies que apresentaram conflitos Mendelianos.

Após a filtragem pelo CQ, ficaram disponíveis os genótipos de 35.521 animais da raça Gir e 28.119 animais da raça Girolando, enquanto os marcadores SNP reduziram para 391.845 na raça Gir e 328.260 na raça Girolando. Os critérios que mais explicaram a diferença numérica pré e pós filtragem foram a frequência mínima de alelos e remoção de progênies em conflitos Mendelianos (remoção de 28.877 e 68.602 SNPs e 31 e 580 animais para as raças Gir e Girolando, respectivamente). Por fim, foi determinada a fase de ligação por meio do software AlphaPhase versão 1.2 (Hickey et al., 2011; Money et al., 2020) cujos comprimentos gerais adotados da cauda e do núcleo à cauda, conforme orientações de Hickey et al. (2011) e Hickey e Money (2020), foram de 400 e 1600 SNPs, respectivamente, para as raças Gir e Girolando.

### **3.2. Testes para identificação de candidatos a haplótipos letais e semiletais**

Seguindo as metodologias propostas por VanRaden et al. (2011a) e Jenko et al. (2019), seis testes distintos foram utilizados para a identificação de candidatos a haplótipos letais e semiletais, sendo agrupados em três categorias: (i) Teste populacional; (ii) Testes de acasalamento de portadores, e; (iii) Testes de semiletalidade.

O Teste populacional e os Testes de acasalamento de portadores baseiam-se na probabilidade de observar ausência completa de animais homozigotos recessivos em toda a população ou nas progênies de acasalamentos portadores, assumindo que os alelos recessivos letais são completamente penetrantes. Os Testes de semiletalidade avaliam a hipótese nula de que o número de animais homozigotos observado e esperado são iguais e o número de animais homozigotos observado difere de zero. Assim, é possível a identificação de alelos semiletais recessivos, com alguns homozigotos não sofrendo efeitos deletérios ou alguns haplótipos marcando imperfeitamente um alelo recessivo letal (Jenko, et al. 2019).

Para melhor compreensão dos cálculos realizados em cada teste, a sigla HH será utilizada para representar animais homocigotos dominantes, Hh para representar animais heterocigotos (portadores) e hh para representar animais homocigotos recessivos para um haplótipo que carrega alelos supostamente letais.

### 3.2.1 Teste populacional

A primeira categoria de testes, proposta por VanRaden et al. (2011a), identificou haplótipos que carregavam alelos recessivos supostamente letais de acordo com a frequência populacional. Com base na frequência de animais portadores ( $f_{Hh}$ ) e do número de animais genotipados (N), a probabilidade de não observar nenhum animal homocigoto ( $P_{hh}$ ) foi calculada por:

$$P_{hh} = [(1 - f_{Hh}^2)/4]^N \quad (1)$$

A esperança do número de animais homocigotos recessivos ( $E_{hh}$ ) é calculada por:

$$E_{hh} = (N/4) \times f_{Hh}^2 \quad (2)$$

### 3.2.2 Testes de acasalamento de portadores

A segunda categoria de testes, proposta por VanRaden et al. (2011a) e adaptada por Jenko et al. (2019), consistiu em identificar haplótipos que carregam alelos recessivos supostamente letais com base no número esperado de progênies homocigotas recessivas de acasalamentos entre animais heterocigotos (Hh×Hh). Aqui foram abordados dois tipos de acasalamentos entre animais heterocigotos: touro ( $n_s$ ) × matriz ( $n_d$ ) e touro ( $n_s$ ) × avô materno ( $n_{mgs}$ ). A probabilidade de não observar nenhuma progênie homocigota recessiva ( $P_{hh}$ ) segue um processo de Bernoulli e foi ser calculada como:

- Para acasalamentos touro × matriz:

$$P_{hh} = 0,75^{n_s \times n_d} \quad (3)$$

- Para acasalamentos touro × avô materno:

$$P_{hh} = 0,875^{n_s \times n_{mgs}} \quad (4)$$

A esperança do número de progêniez homozigotas recessivas ( $E_{hh}$ ) foi calculada como:

- Para acasalamentos touro × matriz:

$$E_{hh} = n_s \times n_d / 4 \quad (5)$$

- Para acasalamentos touro × avô materno:

$$E_{hh} = n_s \times n_{mgs} / 8 \quad (6)$$

### 3.2.3 Testes de semiletalidade

A terceira categoria de testes, proposta por Jenko et al. (2019), avaliou a hipótese nula de que o número observado e esperado de animais não diferiu a partir dos acasalamentos entre animais portadores (touro × matriz ou touro × avô materno). Com base no número de animais homozigotos recessivos observados e esperados ( $O_{hh}$  e  $E_{hh}$ ), heterozigotos ( $O_{Hh}$  e  $E_{Hh}$ ) e homozigotos dominantes ( $O_{HH}$  e  $E_{HH}$ ), foi calculada uma estatística de teste qui-quadrado unidirecional com um grau de liberdade:

$$\chi^2 = (O_{hh} - E_{hh})^2 / E_{hh} + (O_{Hh + HH} - E_{Hh + HH})^2 / E_{Hh + HH} \quad (7)$$

Nesta pesquisa foi estendida para a população a análise de semiletalidade proposta por Jenko et al. (2019), no intuito de observar a ocorrência de haplótipos que carregam alelos recessivos supostamente semiletais dentre os haplótipos analisados pela metodologia de VanRaden et al. (2011a).

### 3.3 Seleção de um haplótipo como recessivo letal ou semiletal

Uma vez que os testes de identificação de candidatos a haplótipos letais e semiletais podem resultar em milhares de haplótipos com ausência completa de animais homozigotos ou alguns homozigotos não sofrendo efeitos deletérios, fez-se necessário adotar critérios para a seleção de um haplótipo como recessivo letal ou para a seleção de um haplótipo como recessivo semiletal.

Neste sentido, a seleção de um candidato a haplótipo recessivo letal foi realizada a partir das condições estabelecidas em Wu et al. (2019), Jenko et al. (2019) e Schmidt et al. (2023): (i) A frequência do haplótipo deveria ser superior a 1%; (ii) A frequência de animais portadores deveria ser superior a 2%; (iii) O número de indivíduos homozigotos observados para o haplótipo deveria ser igual a zero; (iv) O número de indivíduos homozigotos esperados para o haplótipo deveria ser maior ou igual a 1, e; (5) A diferença entre o haplótipo observado e esperado deveria ser significativa após a correção de teste múltiplo de Bonferroni ( $p < 0,05/M$ , onde M é o número de haplótipos com frequência maior que 1%).

Para a seleção de um candidato a haplótipo recessivo semiletal, partindo das condições estabelecidas e readaptadas de Jenko et al. (2019), tem-se: (i) A frequência do haplótipo deveria ser superior a 1%; (ii) A frequência de animais portadores deveria ser superior a 2%; (iii) O número de indivíduos homozigotos esperados para o haplótipo deveria ser maior que o número de indivíduos homozigotos observados; (iv) O número de indivíduos homozigotos observados para o haplótipo deveria ser maior que zero, e; (v) A diferença entre os haplótipos observados e esperados deveria ser significativa após a correção de teste múltiplo de Bonferroni ( $p < 0,05/M$ , onde M é o número de haplótipos com frequência maior que 1%).

Assim, todos os haplótipos que satisfizeram esses critérios foram considerados candidatos a haplótipos letais e/ou semiletais. Nos casos em que houveram centenas de haplótipos candidatos, para filtrar os supostos haplótipos letais ou semiletais diminuindo a taxa de falsas descobertas, foram selecionadas regiões que abrigassem pelo menos cinco haplótipos significativos, chamadas de regiões candidatas, semelhante ao realizado por Jenko et al. (2019). Por fim, um mesmo haplótipo que apresentasse acima de dois testes significativos dentro das respectivas regiões

candidatas foi selecionado como suposto haplótipo letal e/ou semiletal para realização da análise de enriquecimento.

Devido ao fato de ocorrer segregação de alelos das raças Gir e Holandesa para a raça Girolando, regiões de haplótipos com sobreposições foram também analisadas com a finalidade de encontrar possíveis haplótipos letais ou semiletais segregando alelos da raça Gir para a raça Girolando.

### **3.4. Identificação de genes candidatos, análise de enriquecimento funcional, genes ortólogos e fenótipos de letalidade**

Para identificação de genes candidatos dentro dos supostos haplótipos letais e semiletais, os recursos de anotação foram realizados através do software “Bioconductor” versão 3.18 com o pacote “biomaRt” versão Ensembl Genes 103 (Durinck et al., 2009), usando o *btaurus\_gene\_ensembl* e o genoma de referência *Bos taurus* ARS-UCD 1.2 (Rosen et al., 2020), sendo análise de enriquecimento funcional realizada via MeSH, presente no software “Bioconductor” versão 3.18 via pacote “meshes”.

Os genes candidatos registrados na análise de enriquecimento funcional foram confrontados com genes ortólogos por meio do pacote “orthogene” via “BiocManager”, utilizando-se a espécie *Bos taurus* como espécie de entrada e *Mus musculus* como espécie de referência. Por fim, apenas os genes ortólogos confirmados foram consultados no banco de dados *Mouse Genome Informatics* (Blake, et al. 2017), com a finalidade de identificar fenótipos de efeitos letais nos haplótipos supostamente letais ou semiletais selecionados.

## **4. RESULTADOS E DISCUSSÕES**

### **4.1 Faseamento do genótipo e montagem dos haplótipos**

Os 391.845 marcadores SNPs para a raça Gir e os 328.260 marcadores SNP para a raça Girolando foram faseados usando a metodologia de faseamento de longo alcance e imputação de biblioteca de haplótipos (LRPHLI) implementada no software

AlphaPhase versão 1.2 (Hickey et al., 2011; Money et al., 2020). Adotando-se os comprimentos gerais da cauda e do núcleo à cauda em 400 e 1600 SNPs, respectivamente, foram obtidos um total de 967 núcleos para a raça Gir e 808 núcleos para a raça Girolando.

#### 4.2 Candidatos a haplótipos letais e semiletais identificados

Aplicando as cinco condições de seleção de candidatos a haplótipos recessivos letais ou semiletais (Wu et al., 2019; Jenko et al. 2019; Schmidt et al., 2023), foram identificados 2.187 haplótipos na raça Gir ( $p < 0,05/M$ , com  $M = 17.157$ , onde  $M$  é o número de haplótipos com frequência maior que 1%) e 29 haplótipos na raça Girolando ( $p < 0,05/M$ , com  $M = 16.343$ , onde  $M$  é o número de haplótipos com frequência maior que 1%), com ausência significativa ou níveis reduzidos de homozigosidade. O comprimento médio dos haplótipos identificados foi de 393 SNPs e 390 SNPs, para as raças Gir e Girolando, respectivamente.

A Tabela 1 mostra a quantidade de haplótipos com ausência significativa por teste e raça. Na raça Gir, o maior número de haplótipos significativos (1902) foi observado para o teste de semiletalidade populacional (T6), seguido pelos testes de acasalamentos touro×matriz (T4), com 116 haplótipos significativos, e o teste populacional (T1) com 106 registros.

**Tabela 1:** Quantidade de haplótipos com ausência significativa ou nível reduzido de homozigosidade por teste e raça.

<b>Raça/Teste*</b>	<b>T1</b>	<b>T2</b>	<b>T3</b>	<b>T4</b>	<b>T5</b>	<b>T6</b>	<b>Total</b>
<b>Gir</b>	106	24	10	116	29	1902	<b>2187**</b>
<b>Girolando</b>	5	0	0	1	0	23	<b>29***</b>

\* T1: Teste populacional; T2: Teste de acasalamento de portadores entre touro × matriz; T3: Teste de acasalamento de portadores entre touro × avô materno; T4: Teste de semiletalidade em cruzamentos touro × matriz; T5: Teste de semiletalidade em cruzamentos touro × avô materno; T6: Teste de semiletalidade populacional;

\*\* Número total de haplótipos com  $p < 2,91 \times 10^{-6}$  para as cinco categorias de teste na raça Gir;

\*\*\* Número total de haplótipos com  $p < 3,05 \times 10^{-6}$  para as cinco categorias de teste na raça Girolando.

Apenas três testes apresentaram haplótipos significativos para a raça Girolando, sendo o maior número de haplótipos observado para o teste de semiletalidade populacional (T6), com 23 haplótipos significativos, enquanto o teste populacional (T1) e o teste de semiletalidade em cruzamentos touro × matriz (T4) apresentaram cinco e um haplótipos significativos, respectivamente. Os testes de acasalamento de portadores entre touro × matriz (T2), acasalamento de portadores entre touro × avô materno (T3) e teste de semiletalidade em cruzamentos touro × matriz (T4) não apresentaram haplótipo significativo.

Para a raça Gir, com exceção do teste de semiletalidade populacional (T6) que foi aplicado pela primeira vez neste estudo, os resultados para o teste de semiletalidade em cruzamentos touro×matriz (T4) apresentando um maior número de haplótipos significativos foi semelhante ao encontrado por Jenko et. al (2019), onde os mesmos autores informaram que este teste apresentou o maior número de haplótipos significativos para as raças analisadas Aberdeen Angus, Charolês, Hereford, Limousin e Simental. O teste de semiletalidade populacional (T6) ter se apresentado com maior número de haplótipos significativos pode ser explicado pelo fato que este teste se aplica a uma maior quantidade de indivíduos se comparados aos testes de letalidade e semiletalidade para portadores (T2, T3, T4 e T5) e pelos haplótipos apresentarem possivelmente uma penetrância incompleta, causando viabilidade reduzida se comparado ao teste populacional (T1), onde ocorre a ausência completa do indivíduo devido a letalidade do haplótipo.

Em relação à raça Girolando, o número de haplótipos significativos (29) foi próximo ao encontrado por Jenko et. al (2019) para a raça Simental (36), mas com diferenciação que, para a raça Simental, todos os testes apresentaram pelo menos um haplótipo significativo. Também é esperado, semelhante ao descrito anteriormente para raça Gir, que o teste de semiletalidade populacional (T6) apresentasse maior número de haplótipos significativos, visto que este teste se aplica a uma maior quantidade de indivíduos se comparados aos demais testes e pelo fato dos haplótipos apresentarem possivelmente uma penetrância incompleta.

Uma possível explicação para o baixo número de haplótipos observados na raça Girolando se comparado à raça Gir é o fato de ocorrer segregação de alelos das

raças Gir e Holandesa, resultando em uma maior diversidade de gametas recombinantes em relação às raças puras.

#### **4.3 Regiões candidatas e haplótipos selecionados que carregam supostos alelos recessivos letais ou semiletais**

Os 2.187 haplótipos identificados na raça Gir estavam compreendidos em 682 regiões, sendo que destas, 52 regiões abrigaram pelo menos cinco testes de haplótipos com ausência significativa ou nível reduzido de homozigose. Para a raça Girolando, os 29 haplótipos encontrados estavam compreendidos em 28 regiões. Nesse sentido, as 52 regiões para raça Gir e as 28 regiões para a raça Girolando foram consideradas regiões candidatas a abrigarem supostos haplótipos recessivos que carregam alelos letais ou semiletais (Tabela Suplementar S1 e Tabela Suplementar S2).

Algumas regiões candidatas encontradas para a raça Gir foram semelhantes em estudos de identificação de haplótipos letais ou semiletais em outras raças, utilizando-se o genoma de referência *Bos taurus* ARS-UCD 1.2 (Rosen et al., 2020), como foram os casos das regiões 21 a 22 Mb no cromossomo 1 para a raça Nelore (Schmidt et al., 2023), 115 a 116 Mb no cromossomo 2 para a raça Nordic Holsteins (Wu et al., 2019), 51 Mb no cromossomo 3 para as raças Ayrshire (Null et al., 2017) e 28 a 29 Mb no cromossomo 23 para as raças Nordic Red (Wu et al., 2020) e Limousin (Jenko et al., 2019). Em relação a raça Girolando, utilizando-se o genoma de referência *Bos taurus* ARS-UCD 1.2 (Rosen et al., 2020), a região 75 Mb no cromossomo 5 foi igualmente identificada no estudo de Wu et al. (2020) de identificação de haplótipos letais para a raça Nordic Red.

Assim, entre as 52 regiões candidatas para a raça Gir, oito haplótipos foram selecionados como supostos haplótipos letais ou semiletais por apresentarem acima de dois testes significativos para o mesmo haplótipo, enquanto para a raça Girolando, entre as 28 regiões consideradas candidatas, os 29 haplótipos foram selecionados como supostos haplótipos letais ou semiletais, sendo que na região 8 a 12 Mb do cromossomo 15, dois haplótipos distintos foram identificados. Os supostos haplótipos letais ou semiletais foram nomeados usando um código que inclui a raça (GI para Gir

e GR para Girolando), o cromossomo e o número de haplótipos dentro da raça (H para haplótipo), conforme observado na Tabela 2.

**Tabela 2:** Haplótipos identificados carregando supostos alelos recessivos letais ou semiletais para as raças Gir e Girolando, e seus respectivos testes significativos.

Raça	BTA <sup>1</sup>	Nome do Haplótipo <sup>2</sup>	Teste <sup>3</sup>	Região <sup>4</sup> (pb)	
				Base Inicial	Base Final
Gir	1	GI1H1	T3, T4, T6	18.145.688	20.634.675
	1	GI1H2	T2, T5, T6	20.636.739	22.590.942
	5	GI5H3	T2, T3, T6	6.988.664	9.950.705
	7	GI7H4	T4, T5, T6	104.378.126	106.190.534
	13	GI13H5	T4, T5, T6	56.086.678	58.605.230
	15	GI15H6	T4, T5, T6	35.856.252	38.100.601
	18	GI18H7	T4, T5, T6	15.841.673	17.936.313
	23	GI23H8	T3, T4, T6	28.617.272	30.147.569
Girolando	2	GR2H1	T6	4.023.227	6.671.897
	2	GR2H2	T6	107.784.810	111.449.789
	3	GR3H3	T6	6.828.655	8.520.080
	3	GR3H4	T6	8.532.372	11.593.768
	3	GR3H5	T6	112.900.098	115.428.358
	5	GR5H6	T6	72.261.503	74.895.901
	6	GR6H7	T1	107.351.828	109.696.218
	7	GR7H8	T6	94.382.446	96.357.863
	7	GR7H9	T1	96.362.114	99.201.790
	9	GR9H10	T6	80.092.976	83.153.499
	10	GR10H11	T4	25.904.617	28.553.762
	15	GR15H12	T1	3.981.309	6.215.736
	15	GR15H13	T6	6.217.670	8.942.722
	15	GR15H14	T6	8.958.170	12.598.897
	15	GR15H15	T6	8.958.170	12.598.897
	15	GR15H16	T6	12.603.659	15.780.592
	15	GR15H17	T6	24.349.656	27.554.299
	15	GR15H18	T6	33.779.894	36.777.621
	15	GR15H19	T6	36.778.352	39.927.601
	15	GR15H20	T6	39.928.902	43.276.395
	18	GR18H21	T6	5.123.821	6.456.945
	18	GR18H22	T6	25.076.554	26.724.516
	18	GR18H23	T6	41.869.937	43.809.285
	20	GR20H24	T1	11.245.471	13.919.535
	21	GR21H25	T1	48.761.111	51.254.603
	21	GR21H26	T6	51.258.111	53.929.532
	21	GR21H27	T6	55.984.667	58.437.510
22	GR22H28	T6	47.032.709	50.948.130	
27	GR27H29	T6	31.521.050	35.013.681	

<sup>1</sup> *Bos taurus* Autossomo;

<sup>2</sup> Código de nomeação do haplótipo que inclui a raça (GI para Gir e GR para Girolando), o cromossomo e o número de haplótipos dentro da raça (H para haplótipo);

<sup>3</sup> T1: Teste populacional; T2: Teste de acasalamento de portadores entre touro × matriz; T3: Teste de acasalamento de portadores entre touro × avô materno; T4: Teste de

semiletalidade em cruzamentos touro × matriz; T5: Teste de semiletalidade em cruzamentos touro × avô materno; T6: Teste de semiletalidade populacional;  
<sup>4</sup> Genoma de referência *Bos taurus* ARS-UCD 1.2 (Rosen et al., 2020).

Dois haplótipos da raça Gir (GI1H2 e GI23H8) ocorreram em regiões semelhantes em estudos de identificação de haplótipos letais ou semiletais para outras raças, como foram os casos dos haplótipos nas regiões 21 a 22 Mb no cromossomo 1 para a raça Nelore (Schmidt et al., 2023) e 28 a 29 Mb no cromossomo 23 para as raças Nordic Red (Wu et al., 2020) e Limousin (Jenko et al., 2019). Em relação aos haplótipos da raça Girolando, apenas o haplótipo GR5H6 na região 75 Mb do cromossomo 5 ocorreu em uma região semelhante ao estudo de Wu et al. (2020) com bovinos Nordic Red. Uma hipótese plausível para tais ocorrências seria que, a partir do gado selvagem Auroque (*Bos primigenius*), as duas subespécies formadas (*Bos taurus taurus* ou bovino taurino, originada do *Bos primigenius primigenius*; e *Bos taurus indicus* ou zebu, originada do *Bos primigenius namadicus*, de acordo com Loftus et al., 1994) conservaram haplótipos ancestrais que podem controlar certas características e/ou doenças, como mostrado no estudo de Dawkins et al. (1989) com humanos e nos estudos de Williamson et al. (2011) e Lloyd et al. (2014) com bovinos. Nesse sentido, a diversidade de estudos sobre letalidade de haplótipos em bovinos corroboram para a suposta letalidade e/ou semiletalidade de haplótipos presentes nestas regiões.

Em relação aos testes de identificação de ausência ou nível reduzido de homoziguidade, os haplótipos GR6H7, GR7H9, GR15H12, GR20H24 e GR21H25 da raça Girolando apresentaram resultados apenas para testes de letalidade, sendo denominados, portanto, como haplótipos putativamente letais. Os haplótipos GI7H4, GI13H5, GI15H6 e GI18H7 da raça Gir e os demais 24 haplótipos da raça Girolando apresentaram resultados apenas para testes semiletais, sendo denominados, portanto, como haplótipos putativamente semiletais. Por fim, os haplótipos GI1H1, GI1H2, GI5H3 e GI23H8 na raça Gir apresentaram resultados significativos tanto para testes de letalidade quanto para testes de semiletalidade, indicando uma possível penetrância parcial. Reforça-se que no caso da raça Girolando, o fato de mais de 80% dos haplótipos putativamente deletérios apresentarem testes apenas de semiletalidade (penetrância incompleta) pode estar relacionado à segregação de

alelos das raças Gir e Holandesa (maior diversidade de gametas recombinantes). A penetrância incompleta pode ser influenciada por fatores externos ou internos, como ambiente, outros genes (epistasia), ou até mesmo variabilidade aleatória. Esses fatores podem afetar a expressão do haplótipo, fazendo com que a característica associada apareça em alguns indivíduos, mas não em outros.

Tratando-se da raça Gir, ressalta-se que há possibilidade de terem ocorrido testes de letalidade de haplótipos nas regiões onde apresentaram apenas testes de semiletalidade visto que, para diminuir a taxa de falsas descobertas, foram selecionadas apenas regiões que abrigassem pelo menos cinco haplótipos significativos e que um mesmo haplótipo apresentasse acima de dois testes significativos dentro das respectivas regiões. Nesse sentido, possivelmente todas as regiões candidatas na raça Gir apresentam haplótipos tanto para letalidade quanto para semiletalidade nos casos de penetrância incompleta.

As estatísticas dos supostos haplótipos letais e semiletais para as raças Gir e Girolando são apresentadas na Tabela 3. Os oito supostos haplótipos recessivos letais e semiletais para a raça Gir apresentaram uma frequência média de 7,67%, o que representou 1,91 e 1,74 vezes superior as frequências relatadas por Wu et al. (2019) para bovinos Nordic Holstein (4,01%) e Schmidt et al. (2023) para bovinos Nelore (4,41%), enquanto representou 1,29 e 2,25 vezes menor que as frequências relatadas por Wu et al. (2020), para bovinos Danish Jersey, e Jenko et al. (2019) para cinco raças de gado de corte (17,22%). As frequências para os oito haplótipos foram de 6,00% para GI1H1, 10,30% para GI1H2, 5,16% para GI5H3, 8,61% para GI7H4, 6,26% para GI13H5, 8,14% para GI15H6, 7,87% para GI18H7 e 9,05% para GI23H8, sendo a maior frequência observada para o haplótipo GI1H2 e a menor frequência observada para o haplótipo GI5H3.

**Tabela 3:** Estatísticas para os supostos haplótipos letais e semiletais para as raças Gir e Girolando.

Raça	Haplótipo	Frequência do haplótipo (%)	Número de homozigotos recessivos observados	Homozigotos recessivos esperados	Nº acasalamentos touro×matriz	Nº de homozigotos recessivos observados de acasalamentos touro×matriz	Nº de homozigotos recessivos esperado de acasalamentos touro×matriz	Nº acasalamentos touro×avô materno	Nº de homozigotos recessivos observado de acasalamentos touro×avô materno	Nº de homozigotos recessivos esperado de acasalamentos touro×avô materno	Probabilidade de não observar homozigotos recessivos para teste T1 *	Probabilidade de não observar homozigotos recessivos para teste T2 *	Probabilidade de não observar homozigotos recessivos para teste T3 *	Probabilidade de não observar homozigotos recessivos para teste T4 *	Probabilidade de não observar homozigotos recessivos para teste T5 *	Probabilidade de não observar homozigotos recessivos para teste T6 *
Gir	GI1H1	6,00	33	86	318	4	79	163	0	20	-	-	$3,53 \times 10^{-10}$	0	-	$6,80 \times 10^{-9}$
	GI1H2	10,30	9	260	644	0	161	437	1	54	-	$3,46 \times 10^{-81}$	-	-	$8,77 \times 10^{-15}$	0
	GI5H3	5,16	2	62	204	0	51	159	0	19	-	$3,25 \times 10^{-26}$	$6,02 \times 10^{-10}$	-	-	$9 \times 10^{-15}$
	GI7H4	8,61	59	173	562	21	140	306	6	38	-	-	-	0	$2,48 \times 10^{-8}$	0
	GI13H5	6,26	44	97	245	10	61	207	2	25	-	-	-	$3,99 \times 10^{-14}$	$5,21 \times 10^{-7}$	$4,68 \times 10^{-8}$
	GI15H6	8,14	58	166	237	26	59	325	8	40	-	-	-	$6,10 \times 10^{-7}$	$4,46 \times 10^{-8}$	0
	GI18H7	7,87	74	154	587	21	146	637	6	79	-	-	-	0	0	$3,88 \times 10^{-11}$
	GI23H8	9,05	35	176	404	1	101	151	0	18	-	-	$1,75 \times 10^{-9}$	0	-	0
Girolando	GR2H1	3,74	1	24	18	0	4	16	0	2	-	-	-	-	-	$1,36 \times 10^{-6}$
	GR2H2	4,76	11	41	18	2	4	17	0	2	-	-	-	-	-	$1,94 \times 10^{-6}$
	GR3H3	5,17	15	49	6	1	1	7	1	0	-	-	-	-	-	$6,45 \times 10^{-7}$
	GR3H4	7,35	47	95	23	5	5	17	2	2	-	-	-	-	-	$5,77 \times 10^{-7}$

	GR3H5	9,38	100	158	90	14	22	45	5	5	-	-	-	-	-	2,19×10 <sup>-6</sup>
	GR5H6	4,53	6	40	22	2	5	27	0	3	-	-	-	-	-	4,81×10 <sup>-8</sup>
	GR6H7	2,57	0	12	6	0	1	13	0	1	2,47×10 <sup>-6</sup>	-	-	-	-	-
	GR7H8	6,39	15	79	42	8	10	25	0	3	-	-	-	-	-	1,79×10 <sup>-13</sup>
	GR7H9	2,92	0	16	7	0	1	2	0	0	8,10×10 <sup>-8</sup>	-	-	-	-	-
	GR9H10	9,73	72	159	90	7	22	203	13	25	-	-	-	-	-	1,56×10 <sup>-12</sup>
	GR10H11	11,24	157	173	162	10	40	305	13	38	-	-	-	3,13×10 <sup>-8</sup>	-	-
	GR15H12	3,01	0	16	3	0	0	7	0	0	1,11×10 <sup>-7</sup>	-	-	-	-	-
	GR15H13	6,62	34	83	49	9	12	57	7	7	-	-	-	-	-	4,65×10 <sup>-8</sup>
	GR15H14	4,04	4	33	8	2	2	8	0	1	-	-	-	-	-	2,20×10 <sup>-7</sup>
	GR15H15	8,63	79	144	50	10	12	98	11	12	-	-	-	-	-	2,91×10 <sup>-8</sup>
	GR15H16	11,05	118	182	91	16	22	150	16	18	-	-	-	-	-	1,14×10 <sup>-6</sup>
	GR15H17	8,98	84	137	111	15	27	95	6	11	-	-	-	-	-	3,00×10 <sup>-6</sup>
Girolando	GR15H18	7,29	52	106	42	8	10	126	11	15	-	-	-	-	-	6,99×10 <sup>-8</sup>
	GR15H19	5,38	16	56	36	2	9	17	0	2	-	-	-	-	-	5,04×10 <sup>-8</sup>
	GR15H20	5,73	19	59	40	3	10	28	2	3	-	-	-	-	-	1,04×10 <sup>-7</sup>
	GR18H21	6,71	39	88	41	13	10	77	11	9	-	-	-	-	-	9,00×10 <sup>-8</sup>
	GR18H22	5,39	20	58	31	7	7	49	3	6	-	-	-	-	-	3,12×10 <sup>-7</sup>
	GR18H23	5,20	12	43	12	2	3	26	2	3	-	-	-	-	-	1,35×10 <sup>-6</sup>
	GR20H24	3,35	0	19	19	0	4	40	0	5	2,27×10 <sup>-9</sup>	-	-	-	-	-
	GR21H25	2,67	0	13	-	-	-	-	-	-	1,18×10 <sup>-6</sup>	-	-	-	-	-
	GR21H26	6,99	39	91	52	7	13	28	1	3	-	-	-	-	-	2,64×10 <sup>-8</sup>
	GR21H27	6,41	28	74	26	4	6	10	1	1	-	-	-	-	-	4,67×10 <sup>-8</sup>
	GR22H28	3,83	3	27	22	2	5	6	0	0	-	-	-	-	-	2,97×10 <sup>-6</sup>
	GR27H29	4,35	8	36	14	2	3	8	1	1	-	-	-	-	-	2,12×10 <sup>-6</sup>

\* T1: Teste populacional; T2: Teste de acasalamento de portadores entre touro × matriz; T3: Teste de acasalamento de portadores entre touro × avô materno; T4: Teste de semiletalidade em cruzamentos touro × matriz; T5: Teste de semiletalidade em cruzamentos touro × avô materno; T6: Teste de semiletalidade populacional.

Os supostos haplótipos recessivos letais e semiletais para a raça Girolando apresentaram uma frequência média de 5,98%, o que representou 1,49 e 1,36 vezes superior às frequências relatadas por Wu et al. (2019) e Schmidt et al. (2023) para bovinos Nordic Holstein e Nelore, respectivamente, enquanto representou 1,65 e 2,88 vezes menor que as frequências relatadas por Wu et al. (2020) e Jenko et al. (2019) para bovinos Danish Jersey e para cinco raças de gado de corte, respectivamente. A maior frequência foi observada para o haplótipo GR10H11 (11,24%), enquanto a menor foi observada para o haplótipo GR21H25 (2,67%). Em geral, os resultados mostram que o número de homozigotos recessivos observados foi, em maioria, menor que o número esperado para os diferentes tipos de testes.

Para a raça Gir, os haplótipos GI1H1, GI1H2, GI5H3 e GI23H8 apresentaram resultados tanto para testes de letalidade (T2 para GI1H2 e GI5H3; T3 para GI1H1, GI5H3 e GI23H8) quanto para testes de semiletalidade (T4 para GI1H1 e GI23H8; T5 para GI1H2; T6 para GI1H1, GI1H2, GI5H3 e GI23H8), enquanto os haplótipos GI7H4, GI13H5, GI15H6 e GI18H7 apresentaram resultados apenas para testes semiletais (T4, T5 e T6), conforme já descrito anteriormente. Tais características foram parcialmente semelhantes às estatísticas de haplótipos apresentadas por Jenko et al. (2019), que encontraram haplótipos indicando tanto letalidade quanto semiletalidade nos estudos de cinco raças distintas.

O haplótipo GI5H3 apresentou ausência completa de homozigotos em dois testes de letalidade, o que pode indicar uma maior chance de letalidade deste haplótipo em cruzamentos. Os haplótipos GI1H2 e GI23H8, embora tenham apresentado letalidade apenas nos testes T2 e T3, respectivamente, mostraram também uma tendência maior para letalidade em cruzamentos visto que, para os resultados dos testes T2 (GI23H8) e T3(GI1H2), ambos os haplótipos apresentaram apenas um homozigoto recessivo. O mesmo pode-se estender para o haplótipo GI1H1 que apresentou apenas quatro homozigotos recessivos.

Para a raça Girolando, os haplótipos GR6H7, GR7H9, GR15H12, GR20H24 e GR21H25 foram significativos para testes de letalidade, o que é de fato esperado para o haplótipo visto que apresentam ausência completa

de homozigotos observados para os três tipos de testes de letalidade (populacional e acasalamento de portadores). O haplótipo GR2H1, embora tenha se apresentado significativo para teste de semiletalidade, pode-se ser considerado como letal, visto que há apenas um homozigoto observado para o teste populacional. Os demais haplótipos foram significativos apenas para testes de semiletalidade, isto é, não houve haplótipos letais e semiletais, simultaneamente, como já observado anteriormente na Tabela 2.

Buscando encontrar possíveis haplótipos letais ou semiletais que segregam alelos da raça Gir para a raça Girolando, ao realizar a comparação das regiões identificadas neste estudo, foram observadas duas regiões com sobreposições parciais (Tabela Suplementar S3). Estas sobreposições podem abrigar haplótipos em segregação alélica letal ou semiletal, visto que animais da raça Girolando possuem alguma fração da composição racial compartilhada com a raça Gir. As regiões foram nomeadas usando um código que une os códigos de ambas as raças (GIGR, sendo GI e GR representando as raças Gir e Girolando, respectivamente), o cromossomo e o número da região compartilhada (R para região), conforme Tabela 4.

**Tabela 4:** Regiões identificadas carregando supostos haplótipos letais ou semiletais em possível segregação alélica da raça Gir para a raça Girolando.

BTA <sup>1</sup>	Nome da Região <sup>2</sup>	Teste <sup>3</sup>	Região <sup>4</sup> (pb)	
			Base Inicial	Base Final
3	GIGR3R1	T1, T2, T6	10.207.866	11.593.768
15	GIGR15R2	T4, T5, T6	35.856.252	36.777.621

<sup>1</sup> *Bos taurus* Autossomo;

<sup>2</sup> Código de nomeação da região que inclui ambas as raças (GIGR, sendo GI para a raça Gir e GR para a raça Girolando), o cromossomo e o número da região compartilhada (R para região);

<sup>3</sup> T1: Teste populacional; T2: Teste de acasalamento de portadores entre touro × matriz; T3: Teste de acasalamento de portadores entre touro × avô materno; T4: Teste de semiletalidade em cruzamentos touro × matriz; T5: Teste de semiletalidade em cruzamentos touro × avô materno; T6: Teste de semiletalidade populacional;

<sup>4</sup> Genoma de referência *Bos taurus* ARS-UCD 1.2 (ROSEN et al., 2020).

As duas regiões que carregam possíveis alelos em segregação apresentaram haplótipos significativos tanto para testes de letalidade (T1 e T2 para a região GIGR3R1) quanto para testes de semiletalidade (T6 para a região GIGR3R1 e T4, T5 e T6 para a região GIGR15R2).

**Tabela 5:** Estatísticas dos supostos haplótipos letais ou semiletais em segregação entre as raças Gir e Girolando.

Região	Haplótipo	Frequência do haplótipo (%)	Número de homozigotos recessivos observados	Homozigotos recessivos esperados	Nº acasalamentos touro×matriz	Nº de homozigotos recessivos observados de acasalamentos touro×matriz	Nº de homozigotos recessivos esperado de acasalamentos touro×matriz	Nº acasalamentos touro×avô materno	Nº de homozigotos recessivos observado de acasalamentos touro×avô materno	Nº de homozigotos recessivos esperado de acasalamentos touro×avô materno	Probabilidade de não observar homozigotos recessivos para teste T1 *	Probabilidade de não observar homozigotos recessivos para teste T2 *	Probabilidade de não observar homozigotos recessivos para teste T3 *	Probabilidade de não observar homozigotos recessivos para teste T4 *	Probabilidade de não observar homozigotos recessivos para teste T5 *	Probabilidade de não observar homozigotos recessivos para teste T6 *
GIGR3R1	GI3R1H1	5,69	1	85	12	0	3	-	-	-	-	-	-	-	-	0
	GI3R1H2	6,15	42	96	23	4	5	1	0	0	-	-	-	-	-	2,31×10 <sup>-8</sup>
	GI3R1H3	2,98	33	21	74	0	18	40	1	5	-	5,68×10 <sup>-10</sup>	-	-	-	-
	GI3R1H4	2,25	0	13	2	0	0	-	-	-	1,39×10 <sup>-6</sup>	-	-	-	-	-
	GI3R1H5	6,76	35	117	61	12	15	7	1	0	-	-	-	-	-	1,17×10 <sup>-14</sup>
	GR3R1H6	7,35	47	95	23	5	5	17	2	2	-	-	-	-	-	5,77×10 <sup>-7</sup>
GIGR15R2	GI15R2H7	7,36	31	139	33	5	8	1	0	0	-	-	-	-	-	0
	GI15R2H8	5,89	6	90	18	3	4	1	0	0	-	-	-	-	-	0
	GI15R2H9	8,14	58	166	237	26	59	325	8	40	-	-	-	6,10×10 <sup>-7</sup>	4,46×10 <sup>-8</sup>	0
	GR15R2H10	7,29	52	106	42	8	10	126	11	15	-	-	-	-	-	6,99×10 <sup>-8</sup>

\* T1: Teste populacional; T2: Teste de acasalamento de portadores entre touro × matriz; T3: Teste de acasalamento de portadores entre touro × avô materno; T4: Teste de semiletalidade em cruzamentos touro × matriz; T5: Teste de semiletalidade em cruzamentos touro × avô materno; T6: Teste de semiletalidade populacional.

Foram observados um total de 10 haplótipos de ambas as raças (Tabela 5), sendo oito haplótipos da raça Gir e dois haplótipos da raça Girolando. Os haplótipos presentes nas duas regiões foram nomeados usando um código que une os códigos de ambas as raças (GIGR, sendo GI representando a raça Gir e GR representando a raça Girolando), o cromossomo, o número da região compartilhada (R para região) e o número de haplótipos dentro da raça (H para haplótipo).

Os haplótipos na raça Gir em provável segregação alélica apresentaram uma frequência média de 5,65%, enquanto os haplótipos em mesma região, presentes na raça Girolando, apresentaram uma frequência média de 7,36%. O haplótipo GI15R2H9 é o mesmo haplótipo GI15H6 indicado como haplótipo putativamente semiletal para a raça Gir, o que pode reforçar segregação alélica semiletal para a raça Girolando na região do cromossomo 15. O haplótipo GI3R1H4 na raça Gir apresenta número de homocigotos nulo para dois testes de letalidade simultaneamente, enquanto o haplótipo GR3R1H6 na raça Girolando apresenta redução de homocigotos na população, o que pode indicar uma possível segregação alélica semiletal para o Girolando na região do cromossomo três.

Assim, os haplótipos presentes nessas regiões passaram por análise funcional para identificação de genes candidatos que foram, em seguida, submetidos a ortologia para levantamento de possíveis fenótipos de letalidade.

#### **4.4 Genes candidatos, análise de enriquecimento funcional, genes ortólogos e fenótipos de letalidade**

Em relação a raça Gir, considerando um nível de significância de 5%, o teste hipergeométrico para a categoria D (compostos químicos e drogas) identificou um termo de ontologia MeSH (Descrição) significativo ( $p$ -valor  $\approx 0,00054$ ), apontando cinco genes (ID gene) que se enquadraram neste termo, conforme Tabela 6.

**Tabela 6:** Descrição do termo MeSH (categoria D) e parâmetro gerado pelo teste hipergeométrico associado aos haplótipos supostamente letais ou semiletais para raça Gir.

Categoria	ID	Descrição	p-valor	p-ajustado	ID gene
Compostos químicos e drogas	D008565	Proteínas de Membrana	0,000536738	0,045622759	<i>TMPRSS15</i>
					<i>SYT1</i>
					<i>GNAS</i>
					<i>ATP5F1E</i>
					<i>CALCB</i>

O termo de ontologia genética (GO) “Proteínas de Membrana” refere-se a proteínas encontradas em membranas, incluindo membranas celulares e intracelulares, consistindo em dois grupos: proteínas periféricas e proteínas integrais. Essas proteínas incluem a maioria das enzimas associadas a membranas, proteínas antigênicas, proteínas de transporte e receptores de drogas, hormônios e lectinas (DeCS, 2024), que podem influenciar em um desenvolvimento fetal tanto de forma positiva, quanto de forma negativa. Os cinco genes que se enquadraram neste termo (*SYT1*, *GNAS*, *TMPRSS15*, *CALCB*, *ATP5F1E*) são descritos na Tabela Suplementar S4 de acordo com o ID Gene, ID Ensembl Gene, haplótipo, região, descrição e classificação.

A análise de ortologia identificou genes ortólogos para os cinco genes candidatos encontrados na análise de enriquecimento. Os genes ortólogos, ao serem pesquisados no banco de dados *Mouse Genome Informatics*, apresentaram fenótipo de letalidade embrionária ou perinatal de dois genes: *sinaptotagmina 1* (*SYT1*) para o haplótipo G15H3 e *locus do complexo GNAS* (*GNAS*) para o haplótipo G113H5 (Tabela 7).

Tratando-se da raça Girolando, considerando um nível de significância de 5% e alinhando os resultados GO com os objetivos desta pesquisa, o teste hipergeométrico para as categorias A (anatomia), D (compostos químicos e drogas) e G (fenômenos e processos), identificou no total 38 termos de ontologia MeSH (Descrição) significativos, apresentando 89 genes (ID gene) descritos na Tabela Suplementar S5 de acordo com o ID Gene, ID Ensembl Gene, haplótipo, região, descrição e classificação. Alguns desses termos estão intimamente ligados no processo de desenvolvimento fetal, como por exemplo, proteínas de membrana, proteínas de ligação ao DNA, RNA, entre outras.

**Tabela 7:** Fenótipo letal para genes presentes nos supostos haplótipos letais ou semiletais para a raça Gir.

ID Gene ( <i>Bos taurus</i> )	ID Gene Ortólogo <sup>1</sup> ( <i>Mus musculus</i> )	Haplótipo <sup>2</sup>	Região <sup>3</sup> (pb)		Fenótipo Letal <sup>4</sup>
			Base Inicial	Base Final	
SYT1	SYT1	GI5H3	8.645.528	9.048.268	Homozigotos nulos não se alimentam, apresentam transmissão sináptica prejudicada e liberação de neurotransmissores evocados por Ca <sup>2+</sup> e morrem com 48 horas de vida.
GNAS	GNAS	GI13H5	57.465.320	57.519.383	Homozigotos mutantes param de se desenvolver normalmente por volta do dia embrionário 10,5 e morrem no pré-natal.

<sup>1</sup> Ortólogo via “orthogene” utilizando-se a espécie *Bos taurus* como espécie de entrada e *Mus musculus* como espécie de referência;

<sup>2</sup> Código de nomeação do haplótipo que abriga o gene, onde inclui a raça (GI para Gir e GR para Girolando), o cromossomo e o número de haplótipos dentro da raça (H para haplótipo);

<sup>3</sup> Localização do gene conforme genoma de referência *Bos taurus* ARS-UCD 1.2 (ROSEN et al., 2020);

<sup>4</sup> Fenótipo de letalidade embrionária, pré-natal ou perinatal apresentado pelo banco de dados *Mouse Genome Informatics*.

A análise de ortologia identificou genes ortólogos para 79 genes candidatos encontrados na análise de enriquecimento (Tabela Suplementar S6) e que, ao serem pesquisados no banco de dados *Mouse Genome Informatics*, apresentaram fenótipo de letalidade ou semiletalidade embrionária, pré-natal ou perinatal para 11 genes (Tabela 8), sendo eles: *proteína C*, *inativador dos fatores de coagulação Va e VIIIa (PROC)*; *mioglobina (MB)*; *cadeia pesada 9 da miosina (MYH9)*; *pró-proteína convertase subtilisina/kexin tipo 1 (PCSK1)*; *regulador transcricional associado YAP1 (YAP1)*; *angiomotina like 1 (AMOTL1)*; *fator de iniciação da tradução eucariótica 4 gama 2 (EIF4G2)*; *adrenomedulina (ADM)*; *transaminase glutâmico-oxaloacética 2 (GOT2)*; *subunidade reguladora 1 da fosfoinositídeo-3-quinase (PIK3R1)* e *cromogranina A (CHGA)*.

**Tabela 8:** Fenótipo letal para genes presentes nos supostos haplótipos letais ou semiletais para a raça Girolando.

ID Gene ( <i>Bos taurus</i> )	ID Gene Ortólogo <sup>1</sup> ( <i>Mus musculus</i> )	Haplótipo <sup>2</sup>	Região <sup>3</sup> (pb)		Fenótipo Letal <sup>4</sup>
			Base inicial	Base final	
<i>PROC</i>	<i>PROC</i>	GR2H1	5.102.056	5.112.985	A inativação do loco resulta em morte dentro de 24 horas após o nascimento devido à coagulopatia de consumo.
<i>MB</i>	<i>MB</i>	GR5H6	73.814.165	73.824.954	Homozigotos para um alelo nulo apresentam aumento da densidade capilar cardíaca, fluxo coronariano, níveis de hemoglobina e hematócrito. Muitos camundongos homozigotos para outro alelo nulo morrem de insuficiência cardíaca no meio da gestação.
<i>MYH9</i>	<i>MYH9</i>	GR5H6	74.719.956	74.805.905	Camundongos nulos homozigotos apresentam letalidade embrionária.
<i>PCSK1</i>	<i>PCSK1</i>	GR7H8	95.745.639	95.790.112	Homozigotos para um alelo nulo apresentam morte pré e pós-natal, baixo peso, diarreia, hipoglicemia, níveis baixos de insulina e GHRH e falta de níveis maduros de glucagon e ACTH. Homozigotos para outro alelo nulo morrem antes da implantação.
<i>YAP1</i>	<i>YAP1</i>	GR15H13	6.501.783	6.634.647	Embriões homozigotos para uma mutação nula deste gene morrem entre os dias embrionários E9,5 e E10,5 devido à avasculogênese do saco vitelino e à falha de ligação entre o alantoide e o córion.
<i>AMOTL1</i>	<i>AMOTL1</i>	GR15H16	15.542.579	15.735.942	Camundongos heterozigotos para uma mutação missense exibem letalidade pós-natal, enquanto os homozigotos exibem letalidade pré-natal e neonatal sem fenótipos craniofaciais ou cardiovasculares aparentes.
<i>EIF4G2</i>	<i>EIF4G2</i>	GR15H20	41.871.084	41.881.439	Camundongos homozigotos para uma mutação nula direcionada

					não sofrem gastrulação e morrem por volta do dia embrionário E11,5.
ADM	ADM	GR15H20	42.377.322	42.382.803	Homozigotos para uma mutação nula direcionada exibem defeitos cardiovasculares, hyrdops fetalis e letalidade no meio da gestação com defeitos em defeitos placentários, vasos sanguíneos fetais prejudicados e remodelação da artéria espiral materna e diminuição do número de células NK uterinas.
GOT2	GOT2	GR18H22	26.447.740	26.471.481	A homozigosidade para knockout ou certas deleções in-frame ou alelos missense é letal embrionária.
PIK3R1	PIK3R1	GR20H24	11.398.588	11.408.015	Homozigotos para uma mutação nula direcionada exibem letalidade perinatal associada a necrose hepática, ascite quilosa, fibras musculares aumentadas, calcificação do tecido cardíaco e hipoglicemia.
CHGA	CHGA	GR21H27	57.571.750	57.585.040	Homozigotos e heterozigotos para um alelo apresentam hipertensão, níveis anormais de plasma e adrenalina e noradrenalina e, em homozigotos, letalidade embrionária parcial.

<sup>1</sup> Ortólogo via “orthogene” utilizando-se a espécie *Bos taurus* como espécie de entrada e *Mus musculus* como espécie de referência;

<sup>2</sup> Código de nomeação do haplótipo que abriga o gene, onde que inclui a raça (GI para Gir e GR para Girolando), o cromossomo e o número de haplótipos dentro da raça (H para haplótipo);

<sup>3</sup> Localização do gene conforme genoma de referência *Bos taurus* ARS-UCD 1.2 (ROSEN et al., 2020);

<sup>4</sup> Fenótipo de letalidade embrionária, pré-natal ou perinatal apresentado pelo banco de dados *Mouse Genome Informatics*.

A análise funcional dos haplótipos presentes nas regiões de sobreposição identificou 46 genes candidatos, sendo que 45 estavam presentes na região GIGR3R1 e apenas um estava presente na região GIGR15R2. Ao submeter tais genes candidatos à análise de ortologia, apenas 12 genes apresentaram ortologia entre *Bos taurus* e *Mus musculus*. Os genes ortólogos, ao serem pesquisados no banco de dados *Mouse Genome*

*Informatics*, apresentaram fenótipo de letalidade neonatal ou pós-natal de dois genes (Tabela 9), respectivamente: *espectrina alfa, eritrocítica 1 (SPTA1)* para a região GIGR3R1 e *SRY (região determinante de sexo Y) – box 6 (SOX6)* para a região GIGR15R2.

**Tabela 9:** Fenótipo letal para genes presentes nas regiões sobrepostas abrigando haplótipos de segregação alélica letal ou semiletal para a raça Girolando.

ID Gene ( <i>Bos taurus</i> )	ID Gene Ortólogo <sup>1</sup> ( <i>Mus musculus</i> )	Haplótipo <sup>2</sup>	Região <sup>3</sup> (pb)		Fenótipo Letal <sup>4</sup>
			Base Inicial	Base Final	
SPTA1	SPTA1	GI3R1H4	10.996.158	11.088.116	Homozigotos para mutações espontâneas apresentam anemia microcítica, hipocrômica, hemolítica, icterícia e alta mortalidade neonatal.
		GR3R1H6	10.996.957	11.091.232	
SOX6	SOX6	GI3R2H8	35.856.252	36.545.863	Homozigotos para mutações nulas apresentam miopatia cardíaca, bloqueio cardíaco, atraso no crescimento e letalidade pós-natal precoce.
		GR15R2H10	35.860.286	36.529.021	

<sup>1</sup> Ortólogo via “orthogene” utilizando-se a espécie *Bos taurus* como espécie de entrada e *Mus musculus* como espécie de referência;

<sup>2</sup> Código de nomeação do haplótipo que abriga o gene, onde que inclui a raça (GI para Gir e GR para Girolando), o cromossomo e o número de haplótipos dentro da raça (H para haplótipo);

<sup>3</sup> Localização do gene conforme genoma de referência *Bos taurus* ARS-UCD 1.2 (ROSEN et al., 2020);

<sup>4</sup> Fenótipo de letalidade embrionária, pré-natal ou perinatal apresentado pelo banco de dados *Mouse Genome Informatics*.

#### 4.5 Discussões gerais

De forma geral, alelos recessivos letais e semiletais são geralmente raros e específicos da raça, sendo, portanto, difíceis de detectar em uma população, principalmente em casos de mutações recentes, uma vez que os programas de melhoramento animal limitam a endogamia para preservar a variação genética e reduzir a depressão endogâmica. Entretanto, através testes para identificação de candidatos a haplótipos deletérios, propostas por VanRaden et al. (2011a) e Jenko et al. (2019), bem como critérios rigorosos para seleção de haplótipos putativamente deletérios, propostos por Wu et al. (2019), Jenko et al. (2019) e Schmidt et al. (2023), e através da ortologia de

genes homólogos que divergiram por especiação, pôde-se verificar neste estudo a presença de prováveis alelos recessivos letais e semiletais nas populações de bovinos Gir e Girolando.

Um alto número de candidatos a haplótipos deletérios (2.187) foi observado na raça Gir, possivelmente relacionado ao método de inferência de haplótipos implementado no software AlphaPhase que, segundo Hickey et al. (2011), tem por desvantagem o não alinhamento dos núcleos resultantes do particionamento dos cromossomos. Em contrapartida, foi observado um baixo número de candidatos a haplótipos deletérios na raça Girolando (29), possivelmente ligado à segregação de alelos das raças Gir e Holandesa, resultando em uma maior diversidade de gametas recombinantes em relação às raças puras. Tais resultados indicam a necessidade de se realizar, em estudos posteriores, a inferência dos haplótipos através do método que utiliza comprimento variável (que considera o desequilíbrio de ligação para montagem de haplótipos de tamanhos variados), bem como a comparação do número de candidatos a haplótipos deletérios em outras raças sintéticas (confirmando a influência da diversidade de gametas recombinantes no número reduzido de haplótipos).

Um maior número de supostos haplótipos deletérios foi observado em testes de semiletalidade, indicando a penetrância incompleta desses haplótipos. A penetrância incompleta pode ter um impacto significativo na expressão fenotípica de haplótipos semiletais, ou seja, nem todos os indivíduos que carregam esse haplótipo manifestarão os efeitos deletérios associados a ele. Isso significa que alguns indivíduos poderão sobreviver e até reproduzir, mantendo o haplótipo na população, mesmo que ele possa ser potencialmente prejudicial. Em um cenário com penetrância incompleta, a presença de haplótipos semiletais pode levar a uma variedade de resultados fenotípicos dentro de uma população: morte embrionária, pré-natal ou perinatal, ou até a sobrevivência do indivíduo, dependendo de como a penetrância incompleta se manifesta em cada caso. Isso pode influenciar em estudos de identificação de haplótipos letais, já que os efeitos não serão uniformemente visíveis.

Alguns haplótipos (GI1H2 e GI23H8 na raça Gir e GR5H6 na raça Girolando) ocorreram em regiões semelhantes de estudos de identificação de haplótipos letais ou semiletais para outras raças. Uma hipótese plausível, levantada neste estudo para explicar tais ocorrências, é a conservação de haplótipos ancestrais oriundos do gado selvagem Auroque (*Bos primigenius*). Estes haplótipos podem controlar certas características e/ou doenças, como mostrado no estudo de Lloyd et al. (2014) com bovinos, corroborando para a diversidade de pesquisas sobre letalidade de haplótipos em bovinos, indicando suposta letalidade e/ou semiletalidade de haplótipos em determinadas regiões. Ressalta-se ainda a necessidade de se verificar, em estudos posteriores e utilizando-se outro genoma bovino de referência (UMD 3.0, por exemplo), a ocorrência de diferentes regiões semelhantes para as raças Gir e Girolando e outras raças, que não descritas nesta pesquisa, reforçando a possível letalidade e conservação de haplótipos ancestrais oriundos do Auroque.

Na busca de possíveis haplótipos letais ou semiletais em segregação alélica da raça Gir para a raça Girolando, foram observadas duas regiões com sobreposições parciais. Um total de 10 haplótipos (oito da raça Gir e dois da raça Girolando) foram encontrados nessas regiões. O haplótipo GI15R2H9, equivalente ao haplótipo GI15H6, indicado como haplótipo putativamente semiletal para a raça Gir, reforça a possível segregação alélica semiletal para a raça Girolando na região do cromossomo 15. O haplótipo GI3R1H4 na raça Gir apresentou número de homocigotos nulo para dois testes de letalidade simultaneamente, enquanto o haplótipo GR3R1H6 na raça Girolando apresentou redução significativa de homocigotos na população, o que pode indicar outra possível segregação alélica semiletal para o Girolando na região do cromossomo três. Em estudos posteriores pode-se verificar a possível segregação alélica da raça Holandesa para a raça Girolando, o que pode resultar em um maior número de haplótipos em segregação, visto que a raça Holandesa é hoje a pioneira em estudos de haplótipos deletérios, contendo o maior número de haplótipos letais já identificados entre as raças bovinas.

A identificação de genes candidatos, alinhada a análise de enriquecimento e observação de fenótipos letais de genes ortólogos, mostrou-

se uma metodologia alternativa para corroborar a letalidade ou semiletalidade dos haplótipos putativamente deletérios. Embora o método GWAS seja mais comum para a validação de candidatos a haplótipos letais, nem sempre é possível realizar tal análise visto que, em algumas raças bovinas, fenótipos relacionados ao insucesso reprodutivo (possível morte embrionária ou pré-natal) não são reportados para as associações de criadores. Entretanto, alternativas fenotípicas para validação dos haplótipos deletérios são possíveis, como por exemplo prenhez de novilhas (HP), mortalidade pós-natal (PNM) e capacidade de permanência (STAY), analisadas no estudo de Schmidt et al. (2023). Assim, estudos posteriores para confirmação da letalidade ou semiletalidade dos haplótipos identificados nesta pesquisa serão realizados, a depender do levantamento de informações fenotípicas junto às associações de criadores de Gir e Girolando.

Um total de 11 haplótipos putativamente deletérios (dois para a raça Gir e 9 para a raça Girolando) abrigaram 13 genes candidatos (dois para a raça Gir e 11 para a raça Girolando), com ortologia verificada, que apresentaram letalidade segundo o MGD. O gene candidato *SYT1* corroborou para a letalidade e semiletalidade (nos casos possíveis de penetrância incompleta) do haplótipo G15H3 na raça Gir, uma vez que nenhum indivíduo homozigoto para este haplótipo foi observado nos testes de letalidade para acasalamentos touro×matriz e touro×avô materno, bem como o número observado de homozigotos para o teste de letalidade populacional foi quase nulo (apenas dois indivíduos homozigotos). Algo similar pode-se dizer sobre os genes candidatos *PROC* e *PIK3R1* para os respectivos haplótipos GR2H1 (um indivíduo homozigoto observado) e GR20H24 (nenhum indivíduo homozigoto) na raça Girolando. Nesse sentido, pode-se afirmar que o haplótipo GR20H24 e GR2H1 marcam supostos alelos recessivos letais e semiletais, respectivamente.

Embora os haplótipos GR6H7, GR7H9, GR15H12 e GR21H25 apresentaram ausência de homozigotos completa (para todos os testes de letalidade), não foi encontrado genes candidatos que corroborassem com sua letalidade. Entretanto, deve-se ter atenção para tais haplótipos devido seus status nulo de homozigose e considerar que possíveis erros ocorreram

durante a inferência de haplótipos levando a estes resultados, visto que os haplótipos de regiões adjacentes (GR7H8, GR15H13 e GR21H27) apresentaram genes candidatos com fenótipos de letalidade ou semiletalidade.

A segregação alélica da raça Gir para a raça Girolando na região GIGR3R1 é corroborada pelo gene *SPTA1*, pois nesta região o haplótipo GI3R1H4 não apresenta qualquer homozigoto recessivo (letal para Gir), enquanto o haplótipo GR3R1H6 apresenta redução significativa de homozigotos recessivos pelo teste de semiletalidade populacional (semiletal para Girolando). Na região GIGR15R2 também pode estar ocorrendo uma segregação alélica com penetrância incompleta, corroborada pelo gene *SOX6*, visto que os haplótipos GI15R2H7, GI15R2H8, GI15R2H9 e GR15R2H10, presentes nesta região, apresentaram redução significativa de homozigotos recessivos (semiletal para Gir e Girolando). A penetrância incompleta de tais segregações para a raça Girolando pode ser possivelmente explicada pelo fato que estes animais recebem parte alélica de bovinos da raça Holandesa. Nesse sentido, estudos posteriores analisando possíveis segregações haplotípicas da raça Holandesa para a raça Girolando, nestas mesmas regiões, podem reforçar ainda mais a segregação alélica deletéria das regiões identificadas.

De forma ampla, estudos posteriores utilizando GWAS, que identifica, através de métodos matemáticos, associações entre regiões genômicas e traços fenotípicos, tornam-se necessários para confirmar a letalidade ou semiletalidade, bem como a segregação alélica entre as raças, dos haplótipos identificados nesta pesquisa.

## **5. CONCLUSÃO**

No total foram identificados 11 haplótipos putativamente letais ou semiletais abrigando 13 genes candidatos que apresentaram fenótipos deletérios segundo o MGD. Para os haplótipos GI5H3 e GI13H5 da raça Gir, os genes candidatos *SYT1* e *GNAS* se mostraram letais nas fases perinatal e embrionária, respectivamente, através de análise via ortologia genética. Na raça Girolando, 11 genes candidatos (*PROC*, *MB*, *MYH9*, *PCSK1*, *YAP1*,

*AMOTL1*, *EIF4G2*, *ADM*, *GOT2*, *PIK3R1* e *CHGA*) estavam presentes em oito haplótipos (GR2H1, GR5H6, GR7H8, GR15H13, GR15H16, GR15H20, GR18H22, GR20H24 e GR21H27), apresentando letalidade ou semiletalidade em fases embrionária, pré-natal ou perinatal, através de fenótipo, via ortologia genética, no banco de dados *Mouse Genome Informatics*.

Os genes *SYT1* (GI5H3), *PROC* (GR2H1) e *PIK3R1* (GR20H24) corroboraram para a provável semiletalidade e letalidade dos respectivos haplótipos, enquanto os genes *SPTA1* (região GIGR3R1) e *SOX6* (região GIGR3R2) corroboraram para uma possível segregação alélica semiletal.

Nesse sentido, o status de portador de haplótipo letal ou semiletal pode ser usado na seleção de touros, buscando reduzir a frequência de mutações na população evitando acasalamentos de risco. Assim, as descobertas realizadas neste estudo permitirão a implementação de estratégias de acasalamento, baseadas no genoma, para evitar perdas reprodutivas e, conseqüentemente, perdas econômicas.

## 6. REFERÊNCIAS

ADAMS, H. A.; SONSTEGARD, T. S.; VANRADEN, P. M.; NULL, D. J.; VAN TASSELL, C. P.; LEWIN, H. A. Identification of a nonsense mutation in APAF1 that is causal for a decrease in reproductive efficiency in dairy cattle. **Journal of Dairy Science**, v. 99, n. 8, p. 6693–6701, august 2016. URL: <https://doi.org/10.3168/jds.2015-10517>

AGUILAR, I.; MISZTAL, I.; JOHNSON, D. L.; LEGARRA, A.; TSURUTA, S.; LAWLOR, T. J. A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. **Journal of Dairy Science**, v. 93, p. 743-752, february 2010. URL: <https://doi.org/10.3168/jds.2009-2730>

ALBERTINO, L. G. **Frequência alélica da mutação APAF1 em bovinos da raça holandesa (Holstein-Friesian) no Brasil**. Dissertação (Mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Medicina Veterinária e Zootecnia, Botucatu – SP, 2021.

ARAUJO, A. C. **Predição e associação genômica via haplótipos em ruminantes**. Tese (Doutorado) - Universidade Estadual Do Sudoeste Da Bahia (UESB), Programa de Pós-Graduação em Zootecnia, Itapetinga-BA, 2022.

BLAKE, J. A.; EPPIG, J. T.; KADIN, J. A.; RICHARDSON, J. E.; SMITH, C. L.; BULT, C. J.; et al. Mouse Genome Database (MGD) - 2017: community knowledge resource for the laboratory mouse. **Nucleic Acids Research**. v. 45, p. D723–D729, january 2017. URL: <https://doi.org/10.1093/nar/gkw1040>

BEN BRAIEK, M.; FABRE, S., HOZÉ, C.; ASTRUC, J. M.; ROMIEUX C. M. Identification of homozygous haplotypes carrying putative recessive lethal mutations that compromise fertility traits in French Lacaune dairy sheep. **Genetics Selection Evolution**, v. 53, n. 41, may 2021. URL: <https://doi.org/10.1186/s12711-021-00634-1>

BROWNING, S. R.; BROWNING, B. L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. **Am J Hum Genet**. v. 81, n. 5, p. 1084-1097, november 2007. URL: <https://doi.org/10.1086/521987>

CANAZA-CAYO, A. W.; LOPES, P. S.; COBUCI, J. A.; MARTINS, M. F.; SILVA, M. V. G. B. Genetic parameters of milk production and reproduction traits of Girolando cattle in Brazil. **Italian Journal of Animal Science**, v. 17, n. 1, p. 22-30, june 2018. URL: <https://doi.org/10.1080/1828051X.2017.1335180>

CANAZA-CAYO, A. W.; LOPES, P. S.; SILVA, M. V. G.; COBUCI, J. A.; TORRES, R. A.; MARTINS, M. F.; ARBEX, W. A. Estrutura populacional da

raça Girolando. **Ciência Rural**, v. 44, n. 11, p. 2072-2077, novembro 2014. URL: <https://doi.org/10.1590/0103-8478cr20131307>

CDCB – **Council on Dairy Cattle Breeding**. Bowie, Maryland, United States. URL: <https://www.uscdcb.com>. Acesso em: 02 de março de 2021.

CHARLIER, C.; COPPIETERS, W.; ROLLIN, F.; DESMECHT, D.; AGERHOLM, J. S.; CAMBISANO, N.; CARTA, E.; DARDANO, S.; DIVE, M.; FASQUELLE, C.; FRENNET, J.-C.; HANSET, R.; HUBIN, X.; JORGENSEN, C.; KARIM, L.; KENT, M.; HARVEY, K.; PEARCE, B. R.; SIMON, P.; TAMA, N.; NIE, H.; VANDEPUTTE, S.; LIEN, S.; LONGERI, M.; FREDHOLM, M.; HARVEY, R. J.; GEORGES, M.; 2008. Highly effective SNP-based association mapping and management of recessive defects in livestock. **Nature Genetics**, v. 40, n. 4, p. 449–454, april 2008. URL: <https://doi.org/10.1038/ng.96>

CHOI, Y.; CHAN, A. P. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. **Bioinformatics**, v. 31, n. 16, p. 2745-2747, august 2015. URL: <https://doi.org/10.1093/bioinformatics/btv195>

CLAYTON, D. snpStats: SnpMatrix and XSnpmatrix classes and methods. **R package version**, 1.40.0, 2020.

COLE, J. B.; NULL, D. J.; VANRADEN, P. M. Phenotypic and genetic effects of recessive Haplotipos on yield, longevity, and fertility. **Journal of Dairy Science**, v. 99, n. 9, p. 7274–7288, july 2016. URL: <http://dx.doi.org/10.3168/jds.2015-10777>

COOPER, T. A.; WIGGANS, G. R.; NULL, D. J.; HUTCHISON, J. L.; COLE, J. B. Genomic evaluation, breed identification, and discovery of a haplotype affecting fertility for Ayrshire dairy cattle. **Journal of Dairy Science**, v. 97, n. 6, p. 3878-3882, june 2014. URL: <https://doi.org/10.3168/jds.2013-7427>

COOPER, T. A.; WIGGANS, G. R.; VANRADEN, P. M.; HUTCHISON, J. L.; COLE, J. B.; NULL, D. J. Genomic evaluation of Ayrshire dairy cattle and new haplotypes affecting fertility and stillbirth in Holstein, Brown Swiss and Ayrshire breeds. **Amer. Dairy Sci. Assoc. – Amer. Soc. Anim. Sci. Joint Annual Meeting**, Indianapolis, IN, july 9, poster T206, 2013.

CUYABANO, B.C.D.; SU, G.S.; LUND, M.S. Genomic prediction of genetic merit using LD-based haplotypes in the Nordic Holstein population. **BMC Genomics**, v.15, n.1171, december 2014. URL: <https://doi.org/10.1186/1471-2164-15-1171>

**Descritores em Ciências da Saúde:** DeCS. Disponível em: <http://decs.bvsalud.org/>. Acesso em jan. 2024.

DELANEAU, O.; MARCHINI, J.; ZAGURY J. F. (2012) A linear complexity phasing method for thousands of genomes. **Nat. Methods**. v. 9, n. 2, p. 179-181, december 2012. URL: <https://doi.org/10.1038/nmeth.1785>

DEKKERS, J. C. M. Application of Genomics Tools to Animal Breeding. **Current Genomics**, v. 13, n. 3, p. 207–212, may 2012. URL: <https://doi.org/10.2174%2F138920212800543057>

DERKS, M. F. L. **From sequence to phenotype: the impact of deleterious variation in livestock**. PhD thesis, Wageningen University, Netherlands, 2020. URL: <https://doi.org/10.18174/515207>

DERKS, M. F. L.; MEGENS, H-J.; BOSSE, M.; LOPES, M. S.; HARLIZIUS, B.; GROENEN, M. A. M. A systematic survey to identify lethal recessive variation in highly managed pig populations. **BMC Genomics**, v. 18, n. 858, november 2017. URL: <https://doi.org/10.1186/s12864-017-4278-1>

DURINCK, S.; MOREAU, Y.; KASPRZYK, A.; DAVIS, S.; MOOR, B.; BRAZMA, A.; HUBER, W. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. **Bioinformatics**, v. 21, n. 16, p. 3439–3440, august 2005. URL: <https://doi.org/10.1093/bioinformatics/bti525>

FAO - **Food and Agriculture Organization of the United Nations**. FAO STAT - Livestock Primary. Roma, Italy, 2019. URL: <http://www.fao.org/faostat/en/#data/QL>. Acesso em: 01 de março de 2021.

FERENČAKOVIĆ, M.; HAMZIC, E.; GREDLER, B.; CURIK, I.; SÖLKNER, J. Runs of Homozygosity reveal genome-wide Autozygosity in the Austrian Fleckvieh cattle. **Agriculturae Conspectus Scientificus**, v. 76, n. 4, p. 325–329, july 2011.

FERNANDES, A. R.; PEREIRA, C. M. A.; OLIVEIRA, G. R. A.; GOMES, F. C.; OLIVEIRA, J. C.; REIS FILHO, J. C.; MACHADO, C. H. C.; MORAES, C. R.; SILVA, L. C. M.; PANETTO, J. C. C.; SILVA, M. V. G. B.; MACHADO, M. A.; VERNEQUE, R. S.; MARTINS, M. F.; FAZA, D. R. L. R.; XAVIER, M. C. Programa Nacional de Melhoramento do Gir Leiteiro – 11ª prova de pré-seleção de touros-touros pré-selecionados por meio de avaliação genômica. **Embrapa Gado de Leite – Documentos (INFOTECA-E)**, 2020.

FRITZ, S.; HOZE, C.; REBOURS, E.; BARBAT, A.; BIZARD, M.; CHAMBERLAIN, A.; ESCOUFLAIRE, C.; VANDER JAGT, C.; BOUSSAHA, M.; GROHS, C.; ALLAIS-BONNET, A. An initiator codon mutation in SDE2 causes recessive embryonic lethality in Holstein cattle. **Journal of Dairy Science**, v. 101, n. 7, p. 6220-6231, july 2018. URL: <https://doi.org/10.3168/jds.2017-14119>

FRITZ, S.; CAPITAN, A.; DJARI, A.; RODRIGUEZ, S. C.; BARBAT, A.; BAUR, A.; GROHS, C.; WEISS, B.; BOUSSAHA, M.; ESQUERRÉ, D.; KLOPP, C.; ROCHA, D.; BOICHARD, D. Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2. **PLOS ONE**, v. 8, n. 6, e65550, june 2013. URL: <https://doi.org/10.1371/journal.pone.0065550>

GEORGES, M.; CHARLIER, C.; HAYES, B. Harnessing genomic information for livestock improvement. **Nature Reviews Genetics**, v. 20, p. 135–156, march 2019. URL: <https://doi.org/10.1038/s41576-018-0082-2>

GUTIERREZ-REINOSO, M. A.; APONTE, P. M.; GARCIA-HERREROS, M. Genomic Analysis, Progress and Future Perspectives in Dairy Cattle Selection: A Review. **Animals**, v. 11, n. 599, february 2021. URL: <https://doi.org/10.3390/ani11030599>

HÄGGMAN, J.; UIMARI, P. Novel harmful recessive haplotypes for reproductive traits in pigs. **Journal of Animal Breeding and Genetics**, v. 134, n. 2, p. 129-135, april 2017. URL: <https://doi.org/10.1111/jbg.12240>

HAYES, B.; GODDARD, M. E. The distribution of the effects of genes affecting quantitative traits in livestock. **Genetics Selection Evolution**, v. 33, n. 209, p. 209-229, june 2001. URL: <https://doi.org/10.1186/1297-9686-33-3-209>

HESS, M.; DRUET, T.; HESS, A.; GARRICK, D. Fixed-length haplotypes can improve genomic prediction accuracy in an admixed dairy cattle population. **Genetics Selection Evolution**, v. 49, n. 54, july 2017. URL: <https://doi.org/10.1186/s12711-017-0329-y>

HICKEY, J. M.; MONEY, D. **AlphaPhase Documentation**. September 2020.

HICKEY, J. M.; CROSSA, J.; BABU, R.; DE LOS CAMPOS, G. Factors Affecting the Accuracy of Genotype Imputation in Populations from Several Maize Breeding Programs. **Crop Science**, v. 52, n. 2, p. 654-663, march 2012. URL: <https://doi.org/10.2135/cropsci2011.07.0358>

HICKEY, J. M.; KINGHORN, B. P.; TIER, B.; WILSON, J. F.; DUNSTAN, N.; WERF, J. H. J. A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes. **Genetics Selection Evolution**, v. 43, n. 12, march 2011. URL: <https://doi.org/10.1186/1297-9686-43-12>

HOFF, J. L.; DECKER, J. E.; SCHNABEL, R. D.; TAYLOR, J. F. Candidate lethal haplotypes and causal mutations in Angus cattle. **BMC genomics**, v. 18, n. 799, october 2017. URL: <https://doi.org/10.1186/s12864-017-4196-2>

HOWARD, D. M.; PONG-WONG, R.; KNAP, P. W.; WOOLLIAMS, J. A. Use of haplotypes to identify regions harbouring lethal recessive variants in pigs. **Genetics Selection Evolution**, v. 49, n. 57, july 2017. URL: <https://doi.org/10.1186/s12711-017-0332-3>

HUQUN; FUKUYAMA, S.; MORINO, H.; MIYAZAWA, H.; TANAKA, T.; SUZUKI, T.; KOHDA, M.; KAWAKAMI, H.; OKAZAKI, Y.; SEYAMA, K.; HAGIWARA, K. A quantitatively-modeled homozygosity mapping algorithm, qHomozygosityMapping, utilizing whole genome single nucleotide polymorphism genotyping data. **BMC Bioinformatics**, v. 11, supplement 7, n. S5, october 2010. URL: <https://doi.org/10.1186/1471-2105-11-S7-S5>

JENKO, J.; MCCLURE, M. C.; MATTHEWS, D.; MCCLURE, J.; JOHNSON, M.; GORJANC, G.; HICKEY, J. M. Analysis of a large dataset reveals haplotypes carrying putatively recessive lethal and semi-lethal alleles with pleiotropic effects on economically important traits in beef cattle. **Genetics Selection Evolution**, v. 51, n. 1, p. 9, march 2019. URL: <https://doi.org/10.1186/s12711-019-0452-z>

KHATIB, H.; MALTECCA, C.; MONSON, R. L.; SCHUTZKUS, V.; RUTLEDGE, J. J. Monoallelic maternal expression of STAT5A affects embryonic survival in cattle. **BMC Genetics**, v. 10, n. 13, march 2009. URL: <https://doi.org/10.1186/1471-2156-10-13>

LANDER, E. S.; BOTSTEIN, D. Homozygosity mapping: A way to map human recessive traits with the DNA of inbred children. **Science**, v. 236, n. 4808, p. 1567–1570, june 1987. URL: <https://doi.org/10.1126/science.2884728>

LEHNER, B. Genotype to phenotype: lessons from model organisms for human genetics. **Nature Reviews Genetics**, v. 14, p. 168-178, march 2013. URL: <https://doi.org/10.1038/nrg3404>

LLOYD, S. S.; BAYARD, D.; LESTER, S.; WILLIAMSON, J. F.; STEELE, E. J.; DAWKINS, R. L. Ancestral Haplotypes, Quantal Genomics and Healthy Beef. **Proceedings, 10th World Congress of Genetics Applied to Livestock Production**, 2014. URL: [https://www.asas.org/docs/default-source/wcgalp-posters/465\\_paper\\_8652\\_manuscript\\_198\\_0.pdf?sfvrsn=2](https://www.asas.org/docs/default-source/wcgalp-posters/465_paper_8652_manuscript_198_0.pdf?sfvrsn=2)

LOFTUS, E. T.; MACHUGH, D. E.; BRADLEY, D. G.; SHARP, P. M.; CUNNINGHAM, P. Evidence for two independent domestications of cattle. **Proc Natl Acad Sci U.S.A.**, v. 91, n. 7, p. 2757–2761, march 1994. URL: <https://doi.org/10.1073/pnas.91.7.2757>

MAKGAHLELA, M. L.; BANGA, C. B.; NORRIS, D.; DZAMA, K.; NG'AMBI, J. W. Genetic correlations between female fertility and production traits in South African Holstein cattle. **South African Journal of Animal Science**, v. 37, n. 3, p. 180-188, 2007. URL: <https://doi.org/10.4314/sajas.v37i3.4090>

MCCLURE, M. C.; BICKHART, D.; NULL, D.; VANRADEN, P. M.; XU, L.; WIGGANS, G.; LIU, G.; SCHROEDER, S.; GLASSCOCK, J.; ARMSTRONG, J.; COLE, J. B.; VAN TASSELL, C. P.; SONSTEGARD, T. S. Bovine exome se-quence analysis and targeted SNP genotyping of recessive fertility defects BH1, HH2, and HH3 reveal a putative causative mutation in SMC2 for HH3. **PLOS ONE**, v. 9, n. 3, p. e92769, march 2014. URL: <https://doi.org/10.1371/journal.pone.0092769>

MENEZES, G. D. O.; KATSURAGI, R.; SANTOS JUNIOR, E. D.; CARVALHO FILHO, I.; NOBRE, P.; SOUZA JUNIOR, M. D.; GONDO, A. **Avaliação genômica e resultados para características Slick e Dupla Musculatura de touros da raça Senepol**: edição 2022. Campo Grande, MS: Embrapa Gado

de Corte, 2022. PDF (111 p.) : il. color, - (Documentos / Embrapa Gado de Corte, ISSN 1983-974X ; 305).

MESBAH UDDIN, M.; HOZE, C.; MICHOT, P.; BARBAT, A.; LEFEBVRE, R.; BOUSSAHA, M.; SAHANA, G.; FRITZ, S.; BOICHARD, D.; CAPITAN, A. A missense mutation (p. Tyr452Cys) in the CAD gene compromises reproductive success in French Normande cattle. **Journal of Dairy Science**, v. 102, n. 7, p. 6340–6356, july 2019. URL: <https://doi.org/10.3168/jds.2018-16100>

MOROTA, G.; BEISSINGER, T. M.; PEÑAGARICANO, F. MeSH-Informed Enrichment Analysis and MeSH-Guided Semantic Similarity Among Functional Terms and gene products in chicken. **G3**, v. 6, n. 8, p. 2447-2453, august 2016. URL: <https://doi.org/10.1534/g3.116.031096>

MOROTA, G.; PENAGARICANO, F.; PETERSEN, J. L.; CIOBANU, D. C.; TSUYUZAKI, K.; NIKAIIDO, I. An application of MeSH enrichment analysis in livestock. **Animal Genetics**, v. 46, n. 4, p. 381-387, august 2015. URL: <https://doi.org/10.1111/age.12307>

**National Human Genome Research Institute**. Haplotype. Disponível em: <https://www.genome.gov/genetics-glossary/haplotype>. Acesso em fev. 2024.

NG, P. C.; HENIKOFF, S. Predicting amino acid changes that affect protein function. **Nucleic Acids Research**, v. 31, n. 13, p. 3812-3814, july 2003. URL: <https://doi.org/10.1093/nar/gkg509>

NULL, D. J.; HUTCHINSON, J. L.; BICKHART, D. M.; VANRADEN, P. M.; COLE, J. B. Discovery of a haplotype affecting fertility in Ayrshire dairy cattle and identification of a putative causal variant. **Journal of Dairy Science**, v. 100, (Suppl. 2):199, june 2017.

OJANGO, J. M.; POLLOT, G. E. Genetics of milk yield and fertility traits in Holstein-Friesian cattle on large-scale Kenyan farms. **Journal of Animal Science**, v. 79, n. 7, july 2001. URL: <https://doi.org/10.2527/2001.7971742x>

OLIVEIRA JÚNIOR, G. A.; SANTOS, D. J. A.; CESAR, A. S. M.; BOISON, S. A.; VENTURA, R. V.; PEREZ, B. C.; GARCIA, J. F.; FERRAZ, J. B. S.; GARRICK, D. J. Fine mapping of genomic regions associated with female fertility in Nelore beef cattle based on sequence variants from segregating sires. **Journal of Animal Science and Biotechnology**, v. 10, n. 97, december 2019. URL: <https://doi.org/10.1186/s40104-019-0403-0>

OSPINA, A. M. T. **Avaliação genômica em bovinos da raça Gir de Brasil e Colômbia**. Tese (Doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal – SP, 2020.

PANETTO, J. C. C.; SILVA, M. V. G. B.; VERNEQUE, R. S.; MACHADO, M. A.; FERNANDES, A. R.; MARTINS, M. F.; REIS, D. R. L.; ARBEX, W. A.; OLIVEIRA, J. C.; VENTURA, H. T.; PEREIRA, M. A. Programa Nacional de

Melhoramento do Gir Leiteiro - Sumário Brasileiro de Touros – 6ª Avaliação Genômica de Touros - Resultado do Teste de Progênie – Abril 2023. Juiz de Fora: (**Embrapa Gado de Leite. Documentos, 244**), 2023.

PAUSCH, H.; SCHWARZENBACHER, H.; BURGSTALLER, J. FLISIKOWSKI, K. Homozygous haplotype deficiency reveals deleterious mutations compromising reproductive and rearing success in cattle. **BMC Genomics**, v. 16, n. 312, april 2015. URL: <https://doi.org/10.1186/s12864-015-1483-7>

PERIPOLLI, E.; STAFUZZA, N. B.; MUNARI, D. P.; LIMA, A. L. F.; IRGANG, R.; MACHADO, M. A.; PANETTO, J. C. do C.; VENTURA, R. V.; BALDI, F.; SILVA, M. V. G. B. Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle. **BMC Genomics**. URL: <https://doi.org/10.1186/s12864-017-4365-3>

PRYCE, J. E.; HAILE-MARIAM, M.; GODDARD, M. E.; HAYES, B. J. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. **Genetics Selection Evolution**, v. 46, n. 71, november 2014. URL: <https://doi.org/10.1186/s12711-014-0071-7>

PRYCE, J. E.; COFFEY, M. P.; BROTHERSTONE, S. H.; WOOLLIAMS, J. A. Genetic relationships between calving interval and body condition score conditional on milk yield. **Journal of Dairy Science**, v. 85, n. 6, june 2002. URL: [https://doi.org/10.3168/jds.S0022-0302\(02\)74229-X](https://doi.org/10.3168/jds.S0022-0302(02)74229-X)

R CORE TEAM. **R**: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2020. URL: <http://www.R-project.org/>

REIS FILHO, J. C.; LOPES, P. S.; VERNEQUE, R. D. S.; TORRES, R. D. A.; TEODORO, R. L.; & CARNEIRO, P. L. S. Population structure of Brazilian Gyr dairy cattle. **Revista Brasileira de Zootecnia**, v. 39, n. 12, p. 2640-2645, december 2010. URL: <https://doi.org/10.1590/S1516-35982010001200012>

REMINGTON, D. L.; O'MALLEY, D. M. Whole-genome characterization of embryonic stage inbreeding depression in a selfed loblolly pine family. **Genetics**, v. 155, n. 1, p. 337–348, may 2000. URL: <https://genetics.org/content/155/1/337>

RINALDO, A.; BACANU, S. A.; DEVLIN, B.; SONPAR, V.; WASSERMAN, L.; ROEDER, K. Characterization of multilocus linkage disequilibrium. **Genetic Epidemiology**, v. 28, n. 3, p. 193-206, april 2005. URL: <https://doi.org/10.1002/gepi.20056>

ROCHA, D. T.; CARVALHO, G. R.; RESENDE, J. C. Cadeia produtiva do leite no Brasil: produção primária. **Embrapa Gado de Leite – Circular Técnica 123 (INFOTECA-E)**, 2020.

ROSEN, B. D.; BICKHART, D. M.; SCHNABEL, R. D.; KOREN, S.; ELSIK, C. G.; TSENG, E.; ROWAN, T. N.; LOW, W. Y.; ZIMIN, A.; COULDREY, C.;

HALL, R.; LI, W.; RHIE, A.; GHURYE, J.; MCKAY, S. D.; THIBAUD-NISSEN, F.; HOFFMAN, J.; MURDOCH, B. M.; SNELLING, W. M.; MCDANELD, T. G.; HAMMOND, J. A.; SCHWARTZ, J. C.; NANDOLO, W.; HAGEN, D. E.; DREISCHER, C.; SCHULTHEISS, S. J.; SCHROEDER, S. G.; PHILLIPPY, A. M.; COLE, J. B.; VAN TASSELL, C. P.; LIU, G.; SMITH, T. P. L.; MEDRANO, J. F. De novo assembly of the cattle reference genome with single-molecule sequencing. **GigaScience**, v. 9, n. 3, p. g1aa021, march 2020. URL: <https://doi.org/10.1093/gigascience/g1aa021>

SAHANA, G.; NIELSEN, U. S.; AAMAND, G. P.; LUND, M. S.; GULDBRANDTSEN, B. Novel harmful recessive haplotypes identified for fertility traits in Nordic Holstein cattle. **PLOS ONE**, v. 8, n. 12, p. e82909, december 2013. URL: <https://doi.org/10.1371/journal.pone.0082909>

SANTANA JÚNIOR, M. L.; PEREIRA, R. J.; BIGNARDI, A. B.; AYRES, D. R.; MENEZES, G. R. O.; SILVA, L. O. C.; LEROY, G.; MACHADO, C. H. C.; JOSAHKIAN, L. A.; ALBUQUERQUE, L. G. Structure and genetic diversity of Brazilian Zebu cattle breeds assessed by pedigree analysis. **Livestock Science**, v. 187, may 2016. URL: <https://doi.org/10.1016/j.livsci.2016.02.002>

SARGOLZAEI, M.; CHESNAIS, J. P.; SCHENKEL, F. S. A new approach for efficient genotype imputation using information from relatives. **BMC Genomics**, v. 15, n. 478, june 2014. URL: <https://doi.org/10.1186/1471-2164-15-478>

SAS, I. SAS® 9.4. **Cary, NC, USA: SAS Institute Inc.**, 2013.

SCHMIDT, P. I.; MOTA, L. F. M.; FONSECA, L. F. S.; et al. Identification of candidate lethal haplotypes and genomic association with post-natal mortality and reproductive traits in Nellore cattle. **Scientific Reports**, v. 13, n. 10399, june 2023. URL: <https://doi.org/10.1038/s41598-023-37586-z>

SCHMIDT, P. I. **Identification of candidate lethal haplotypes and recombination events in Nellore cattle**. Dissertação (Mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal – SP, 2020.

SCHÜTZ, E.; WEHRHAHN, C.; WANJEK, M.; BORTFELD, R.; WEMHEUER, W. E.; BECK, J.; BREINIG, B. The Holstein Friesian lethal hap-lotype 5 (HH5) results from a complete deletion of TBF1M and cholesterol deficiency (CDH) from an ERV-(LTR) insertion into the coding region of APOB. **PLOS ONE**, v. 11, n. 6, p. e0154602, april 2016. URL: <https://doi.org/10.1371/journal.pone.0157618>

SCHWARZENBACHER, H.; BURGSTALLER, J.; SEEFRIED, F. R.; WURMSER, C.; HILBE, M.; JUNG, S.; FUERST, C.; DINHOPL, N.; WEISSENBOCK, H.; FUERSTWALTL, B.; DOLEZAL, M.; WINKLER, R.; GRUETER, O.; BLEUL, U.; WITTEK, T.; FRIES, R.; PAUSCH, H. A missense mutation in TUBD1 is associated with high juvenile mortality in Braunvieh and

Fleckvieh cattle. **BMC Genomics**, v. 17, n. 400, may 2016. URL: <https://doi.org/10.1186/s12864-016-2742-y>

SEWALEM, A.; KISTEMAKER, G. J.; MIGLIOR, F. Relationship between female fertility and production traits in Canadian Holsteins. **Journal of Dairy Science**, v. 93, n. 9, p. 4427-4434, september 2010. URL: <https://doi.org/10.3168/jds.2009-2915>

SHANNON, P.; MARKIEL, A.; OZIER, O.; BALIGA, N. S.; WANG, J. T.; RAMAGE, D.; AMIN, N.; SCHWIKOWSKI, B.; IDEKER, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. **Genome Research**, v. 13, n. 11, p. 2498-2504, november 2003. URL: <https://doi.org/10.1101/gr.1239303>

SILVA, M. V. G. B.; MARTINS, M. F.; GONÇALVES, G. S.; PANETTO, J. C. C.; PIVA, L. C.; MACHADO, M. A.; FAZA, D. R. L. R; FERREIRA JÚNIOR, E. Programa de Melhoramento Genético da Raça Girolando - Sumário de Touros - Resultado do Teste de Progênie (Avaliação Genética / Genômica) - junho/2020. Juiz de Fora: (**Embrapa Gado de Leite. Documentos, 248**), 2020.

SOLLERO, B. P.; GRYNBERG, P. Tutorial para análise funcional a partir de estudos de associação genômica ampla e transcriptômicos utilizando o banco de dados MeSH (Medical Subject Headings) no programa R - outubro/2020. Bagé: (**Embrapa Pecuária Sul. Documentos, 165**), 2020.

SONSTEGARD, T. S.; COLE, J. B.; VANRADEN, P. M.; VAN TASSELL, C. P.; NULL, D. J.; SCHROEDER, S. G.; BICKHART, D.; MCCLURE, M. C. Identification of a nonsense mutation in CWC15 associated with decreased reproductive efficiency in Jersey cattle. **PLOS ONE**, v. 8, n. 1, p. e54872, january 2013. URL: <https://doi.org/10.1371/journal.pone.0054872>

**STgenetics**. Disponível em: <https://www.stgen.com>. Acesso em fev. 2024.

STOFFEL, M. A.; JOHNSTON, S. E.; PILKINGTON, J. G.; PEMBERTON, J. M. Purifying and balancing selection on embryonic semi-lethal haplotypes in a wild mammal. **Evolution Letters**, qrad053, october 2023. URL: <https://doi.org/10.1093/evlett/qrad053>

VANRADEN, P. M.; NULL, D.; HUTCHISON, J.; BICKHART, D.; SCHROEDER, S. Jersey haplotype 2 (JH2). Changes to evaluation system (August 2014). **Council on Dairy Cattle Breeding**, 2014.

VANRADEN, P. M.; NULL, D.; HUTCHISON, J.; COOPER, T. New fertility and stillbirth haplotypes and changes in haplotype status. Changes to evaluation system (August 2013). **Council on Dairy Cattle Breeding**, 2013.

VANRADEN, P. M.; OLSON, K. M.; NULL, D. J.; HUTCHISON, J. L. Harmful recessive effects on fertility detected by absence of homozygous haplotypes.

**Journal of Dairy Science**, v. 94, n. 12, p. 6153-6161, december 2011a. URL: <https://doi.org/10.3168/jds.2011-4624>

VANRADEN, P. M., D. J. NULL, K. M. OLSON, AND J. L. HUTCHISON. Reporting of haplotypes with recessive effects on fertility. **Intersire Sireetin**, n. 44, august 2011b. URL: <https://journal.intersire.org/index.php/ib/article/view/1202>. Acesso em: 02 de março de 2021.

VANRADEN, P. M.; O'Connell, J.R.; Wiggans, G.R.; Weigel, K. A. Genomic evaluations with many more genotypes. **Genet. Sel. Evol.**, v. 43 n. 10, march 2011c. URL: <https://doi.org/10.1186/1297-9686-43-10>

VILLUMSEN, T. M.; JANSS, L.; LUND, M. The importance of haplotype length and heritability using genomic selection in dairy cattle. **Journal of Animal Breeding and Genetics**, v.126, n. 1, p. 3-13, february 2009. URL: <https://doi.org/10.1111/j.1439-0388.2008.00747.x>

VIZONÁ, R. G. **Estudo genético de características morfológicas e de produção de embriões in vitro em bovinos da raça Gir leiteiro**. Dissertação (Mestrado) – Universidade de São Paulo. Faculdade de Medicina Veterinária e Zootecnia. Departamento de Nutrição e Produção Animal, Pirassununga – SP, 2020.

WARR, A.; AFFARA, N.; AKEN, B.; BEIKI, H.; BICKHART, D. M.; BILLIS, K.; CHOW, W.; EORY, L.; FINLAYSON, H. A.; FLICEK, P.; et al. An improved pig reference genome sequence to enable pig genetics and genomics research. **Giga Science**, v. 9, n. 6, june 2020. URL: <https://doi.org/10.1093/gigascience/giaa051>

WARREN, W. C.; HILLIER, L. W.; TOMLINSON, C.; MINX, P.; KREMITZKI, M.; GRAVES, T.; MARKOVIC, C.; BOUK, N.; PRUITT, K. D.; THIBAUD-NISSEN, F. et al. A New Chicken Genome Assembly Provides Insight into Avian Genome Structure. **G3: Genes, Genomes, Genetics**, v. 7, n. 1, p. 109-117, january 2017. URL: <https://doi.org/10.1534/g3.116.035923>

WILLIAMSON, J. F.; STEELE, E. J.; LESTER, S.; KALAI, O.; MILLMAN, J. A.; WOLRIGE, L.; BAYARD, D.; MCLURE, C.; DAWKINS, R. L. Genomic evolution in domestic cattle: Ancestral haplotypes and healthy beef. **Genomics**, v. 97, n. 5, p. 304-312, may 2011. URL: <https://doi.org/10.1016/j.ygeno.2011.02.006>

WRIGHT, S. Mendelian analysis of the pure breeds of livestock: I. The measurement of inbreeding and relationship. **Journal of Heredity**, v. 14, n. 8, p. 339–348, november 1923. URL: <https://doi.org/10.1093/oxfordjournals.jhered.a102354>

WU, X.; MESBAH-UDDIN, M.; GULDBRANDTSEN, B.; LUND, M. S.; SAHANA, G. Novel haplotypes responsible for prenatal death in Nordic Red and Danish Jersey cattle. **Journal of Dairy Science**, v. 103, n. 5, p. 4570–4578, may 2020. URL: <https://doi.org/10.3168/jds.2019-17831>

WU, X.; MESBAH-UDDIN, M.; GULDBRANDTSEN, B.; LUND, M. S.; SAHANA, G. Haplotypes responsible for early embryonic lethality detected in Nordic Holsteins. **Journal of Dairy Science**, v. 102, n. 12, p. 11116-11123, september 2019. URL: <https://doi.org/10.3168/jds.2019-16651>

ZADRA, L. F. Seleção de bovinos e interpretação de DEP (Diferença Esperada na Progenie). **Pesquisa & Tecnologia**, v. 9, n. 1, jan-jun 2012. URL: [https://www.agricultura.sp.gov.br/documents/1007647/0/11.%20SELE%C3%87%C3%83O%20DE%20BOVINOS%20E%20INTERPRETA%C3%87%C3%83O%20DE%20DEP%20\(DIFEREN%C3%87A%20ESPERADA%20NA%20PROG%C3%80ANIE\)ISSN.pdf/f1821fbe-d2f5-148e-f006-f26399af7caf](https://www.agricultura.sp.gov.br/documents/1007647/0/11.%20SELE%C3%87%C3%83O%20DE%20BOVINOS%20E%20INTERPRETA%C3%87%C3%83O%20DE%20DEP%20(DIFEREN%C3%87A%20ESPERADA%20NA%20PROG%C3%80ANIE)ISSN.pdf/f1821fbe-d2f5-148e-f006-f26399af7caf)

## **TABELAS SUPLEMENTARES**

**Tabela Suplementar S1:** Lista de regiões candidatas apresentando pelo menos cinco haplótipos com deficiência homocigótica para a raça Gir.

BTA <sup>1</sup>	Região (pb) <sup>2</sup>		Número de haplótipos significativos <sup>3</sup>
	Base Inicial	Base Final	
1	18.145.688	20.634.675	5
	20.636.739	22.590.942	6
	42.748.345	45.052.090	5
	115.165.794	118.490.366	5
2	4.423.041	6.325.168	5
	82.347.701	86.152.903	6
	114.034.771	116.515.561	5
	132.521.508	134.364.438	5
3	10.207.866	12.773.904	5
	50.088.551	52.847.933	5
4	21.057.855	24.341.716	5
	77.032.763	79.019.142	6
	106.897.626	108.948.065	5
5	6.988.664	9.950.705	7
	27.839.649	32.233.226	5
6	10.206.309	12.242.863	5
7	72.422.640	74.476.430	5
	104.378.126	106.190.534	5
8	6.337.198	8.560.601	5
	9.996.060	10.001.969	6
	14.652.026	16.746.616	5
	90.554.689	92.507.701	5
	99.994.873	100.002.885	5
9	20.692.547	22.671.865	5
	28.575.041	30.647.456	7
10	70.953.425	74.057.399	6
12	9.995.007	10.003.340	5
	10.477.955	12.043.657	5
	49.123.661	51.935.765	5
13	56.086.678	58.605.230	6
14	43.496.343	45.951.267	6
15	31.234.957	33.285.457	5
	35.856.252	38.100.601	5
	53.092.786	55.583.948	5
	66.988.976	69.543.047	5
16	16.067.472	18.264.527	6
	57.641.023	59.827.441	7
	68.448.327	70.407.311	5
	75.275.624	76.690.800	5
17	15.547.341	17.966.225	5
	17.974.044	20.339.693	5
18	15.841.673	17.936.313	7
	20.444.976	22.687.235	5
	57.044.275	61.109.005	5
19	39.014.955	44.115.375	5
20	36.423.134	39.247.015	5
23	28.617.272	30.147.569	6

24	57.346.887	59.136.678	5
	59.141.340	61.001.528	5
25	12.520.823	15.467.899	5
26	3.524.161	5.762.056	5
29	35.670.763	39.764.624	5

<sup>1</sup> Bos taurus Autosomo;

<sup>2</sup> Genoma de referência Bos taurus ARS-UCD 1.2 (ROSEN et al., 2020);

<sup>3</sup> Número de haplótipos com ausência significativa ou nível reduzido de homozigosidade dentro da região.

**Tabela Suplementar S2:** Lista de regiões candidatas apresentando haplótipos com deficiência homocigótica para a raça Girolando.

BTA <sup>1</sup>	Região (pb) <sup>2</sup>		Número de Haplótipos Significativos <sup>3</sup>
	Base Inicial	Base Final	
2	4.023.227	6.671.897	1
	107.784.810	111.449.789	1
3	6.828.655	8.520.080	1
	8.532.372	11.593.768	1
	112.900.098	115.428.358	1
5	72.261.503	74.895.901	1
6	107.351.828	109.696.218	1
7	94.382.446	96.357.863	1
	96.362.114	99.201.790	1
9	80.092.976	83.153.499	1
10	25.904.617	28.553.762	1
15	3.981.309	6.215.736	1
	6.217.670	8.942.722	1
	8.958.170	12.598.897	2
	12.603.659	15.780.592	1
	24.349.656	27.554.299	1
	33.779.894	36.777.621	1
	36.778.352	39.927.601	1
39.928.902	43.276.395	1	
18	5.123.821	6.456.945	1
	25.076.554	26.724.516	1
	41.869.937	43.809.285	1
20	11.245.471	13.919.535	1
21	48.761.111	51.254.603	1
	51.258.111	53.929.532	1
	55.984.667	58.437.510	1
22	47.032.709	50.948.130	1
27	31.521.050	35.013.681	1

<sup>1</sup> Bos taurus Autossomo;

<sup>2</sup> Genoma de referência Bos taurus ARS-UCD 1.2 (ROSEN et al., 2020);

<sup>3</sup> Número de haplótipos com ausência significativa ou nível reduzido de homocigosidade dentro da região.

**Tabela Suplementar S3:** Lista de regiões que possivelmente abrigam haplótipos supostamente letais ou semiletais em segregação da raça Gir para a raça Girolando.

BTA <sup>1</sup>	Região (pb) <sup>2</sup>		Raça <sup>3</sup>
	Base Inicial	Base Final	
3	8.998.750	11.784.663	Girolando
	10.207.866	12.773.904	Gir
15	33.779.894	36.777.621	Girolando
	35.856.252	38.100.601	Gir

<sup>1</sup> Bos taurus Autosomo;

<sup>2</sup> Genoma de referência Bos taurus ARS-UCD 1.2 (ROSEN et al., 2020);

<sup>3</sup> Raça para a qual a região pertence.

**Tabela Suplementar S4:** Genes candidatos encontrados via recurso de anotação e análise de enriquecimento MeSH para haplótipos da raça Gir.

ID Gene <sup>1</sup>	ID Gene Ensembl <sup>2</sup>	Haplótipo <sup>3</sup>	Região (pb) <sup>4</sup>		Descrição <sup>2</sup>	Classificação <sup>2</sup>
			Base inicial	Base final		
<i>TMPRSS15</i>	ENSBTAG00000000597	GI1H1	18.567.865	18.710.975	Serina protease transmembrana 15	
<i>SYT1</i>	ENSBTAG00000034693	GI5H3	8.645.528	9.048.268	Sinaptotagmina 1	
<i>GNAS</i>	ENSBTAG00000052413	GI13H5	57.465.320	57.519.383	Locus do complexo GNAS	Gene codificador de proteína
<i>ATP5F1E</i>	ENSBTAG00000039208	GI13H5	57.357.618	57.361.337	Subunidade épsilon da ATP sintase F1	
<i>CALCB</i>	ENSBTAG00000001516	GI15H6	37806147	37809438	Polipeptídeo beta relacionado à calcitonina	

<sup>1</sup> ID dos genes candidatos encontrados via MeSH;

<sup>2</sup> ID, descrição e classificação encontrados via Ensembl Genes 103 (DURINCK et al., 2009), usando o *btaurus\_gene\_ensembl*;

<sup>3</sup> Código de nomeação do haplótipo que inclui a raça (GI para Gir e GR para Girolando), o cromossomo e o número de haplótipos dentro da raça (H para haplótipo);

<sup>4</sup> Genoma de referência *Bos taurus* ARS-UCD 1.2 (ROSEN et al., 2020).

**Tabela Suplementar S5:** Descrição do termo MeSH (categoria D) e parâmetro gerado pelo teste hipergeométrico associado aos haplótipos supostamente letais ou semiletais para raça Girolando.

<b>Categoria</b>	<b>ID</b>	<b>Descrição</b>	<b>p-valor</b>	<b>p-ajustado</b>	<b>ID gene</b>
Anatomia	D005347	Fibroblasto	$1,14 \times 10^{-7}$	$1,96 \times 10^{-5}$	<i>PTH, RNASE1, MSTN, MMP1, MMP3, HYAL2, MMP13, MYOD1, PIK3R1, MMP15, MMP7, RHOA, LYVE1, SWAP70, FBLN5</i>
	D012259	Ribonuclease Pancreática	$9,53 \times 10^{-11}$	$5,23 \times 10^{-8}$	<i>BRB, RNASE1, RNASE13, ANG, ANG2, RNASE10</i>
	D019714	Inibidor Tecidual de Metaloproteinases	$3,56 \times 10^{-9}$	$1,95 \times 10^{-6}$	<i>MMP1, MMP3, STAR, MMP13, MMP15, MMP7, ADAMTS4</i>
	D008666	Metaloendopeptidases	$7,10 \times 10^{-9}$	$3,90 \times 10^{-6}$	<i>MMP1, MMP3, ADAM2, ADAM9, MMP13, MMP20, MMP15, MMP7, ADAMTS4, NLN</i>
	D002118	Cálcio	$1,92 \times 10^{-8}$	$1,05 \times 10^{-5}$	<i>ADM, PTH, SAG, CAST, CHGA, PROC, SPP2, STAR, CNGB1, MMP13, MMP20, IRAG1, RPGRIP1, RHOA, TRPC6, CACNA1D, PNP, TNNC1, FBLN5, ATP1A4</i>
Compostos químicos e drogas	D007375	Interleucina-1	$3,96 \times 10^{-7}$	0,0002	<i>BRB, SAG, MMP1, MMP3, ADAM9, MMP13, ADAMTS4</i>
	D019717	Inibidor Tecidual da Metaloproteinase-3	$7,39 \times 10^{-7}$	0,0004	<i>MMP1, MMP3, MMP13, MMP15, MMP7, ADAMTS4</i>
	D019716	Inibidor Tecidual da Metaloproteinase-2	$9,73 \times 10^{-7}$	0,0005	<i>LGMN, MMP1, MMP3, MMP13, MMP15, MMP7</i>
	D008565	Proteínas de Membrana	$1,03 \times 10^{-6}$	0,0006	<i>SAG, CHGA, F11R, MMP3, ADAM9, GNAT1, HSPA8, MMP13, IRAG1, MYOD1, PCSK1, RGS9BP, NDUFB1, RHOA, LYVE1, CALCB, ATP1A2, MPZ, AMOTL1, COPA, PGR</i>
	D004791	Inibidores Enzimáticos	$1,12 \times 10^{-6}$	0,0006	<i>ADM, RNASE1, CHGA, GPX1, MMP13, PIK3R1, ADAMTS4, NDUFS2, RHOA, PNP, PRKCD,</i>

				<i>HMOX1, LNPEP, ATP1A4</i>
D020781	Metaloproteinase de matriz 1	1,26×10 <sup>-6</sup>	0,0007	<i>MMP1, MMP3, MMP13, MMP15, MMP7, MMP8</i>
D012260	Ribonucleases	1,68×10 <sup>-6</sup>	0,0009	<i>BRB, RNASE1, RNASE6, RNASE4, ANG</i>
D053504	Metaloproteinases de Matriz Associadas à Membrana	1,68×10 <sup>-6</sup>	0,0009	<i>MMP1, MMP3, MMP13, MMP15, MMP7</i>
D020778	Metaloproteinase de matriz 2	2,06×10 <sup>-6</sup>	0,001	<i>MMP1, MMP3, MMP13, MMP15, MMP7, MMP8</i>
D019715	Inibidor Tecidual da Metaloproteinase-1	4,63×10 <sup>-6</sup>	0,003	<i>MMP1, MMP3, MMP13, MMP15, MMP7</i>
				<i>ADM, PTH, RNASE1, CHGA, MSTN, GNAT1, HSPA8, MMP13, PIK3R1, SPON1, ADAMTS4, GOT2, MYH9, LYVE1, CALCB, DDR2</i>
D010455	Peptídeos	9,16×10 <sup>-6</sup>	0,005	<i>PROC, HSPA8, PIK3R1, NDUFS2, COPB1, COPA</i>
D001224	Ácido Aspártico	1,43×10 <sup>-5</sup>	0,008	<i>MB, PTH, RNASE1, CHGA, GPX1, GNAT1, HYAL2, RNASE6, RHOA, TNNC1, APEH, GLRX, ANG</i>
D014357	Tripsina	1,76×10 <sup>-5</sup>	0,01	<i>MMP3, ADAM2, ADAM9, MMP13, ADAMTS4</i>
D051722	Proteínas ADAM	2,14×10 <sup>-5</sup>	0,012	<i>PTH, GPX1, STAR, CEBPA, HSPA8, MYOD1, PIK3R1, TLR9, USF1, NUCB2, ATF6, SWAP70, ZNF219, POLR2D, YAP1</i>
D004268	Proteínas de Ligação ao DNA	2,32×10 <sup>-5</sup>	0,013	<i>MMP1, MMP3, MMP13, MMP15, MMP7, ADAMTS4</i>
D011480	Inibidores de Protease	2,64×10 <sup>-5</sup>	0,015	<i>RNASE1, CAST, GPR68, MMP20, RNASE6, GOT2, PNP, TNNC1, APEH, SH3BP4, ATP1A4, MANF, ANG</i>
D000596	Aminoácidos	2,70×10 <sup>-5</sup>	0,015	<i>MMP3, ADAM9, MMP13, MMP20</i>
D020782	Metaloproteinases de Matriz	3,93×10 <sup>-5</sup>	0,022	<i>GPX1, SLC12A3, RRAS2, LNPEP, KCNJ11, RAB28, ERLIN2, ATXN3</i>
D024661	Canais de potássio, retificadores internos	5,02×10 <sup>-5</sup>	0,028	

	D017364	Colagenases	$5,30 \times 10^{-5}$	0,029	MMP1, MMP3, ADAM9, MMP13
	D054707	Fosfodiesterases de Nucleotídeos Cíclicos, Tipo 6	$5,88 \times 10^{-5}$	0,032	GPX1, CNGB1, GNAT1, RRAS2, LNPEP, RAB28, ERLIN2, ATXN3
	D011993	Proteínas de Fusão Recombinantes	$6,51 \times 10^{-5}$	0,036	SAG, RNASE1, CHGA, DAG1, GNAI2, GNAT1, HSPA8, IRAG1, MYOD1, PIK3R1, CEP57, TNNC1, POLR2D, COPA
	D012313	RNA	$6,51 \times 10^{-5}$	0,036	PTH, SAG, MSTN, MMP1, MMP3, CEBPA, GNAT1, MMP13, PIK3R1, MMP15, MMP7, EIF4G2, ITLN2
	D053509	Metaloproteinase de Matriz 13	$6,98 \times 10^{-5}$	0,038	MMP1, MMP3, ADAM9, MMP13
	D046911	Substâncias Macromoleculares	$7,69 \times 10^{-5}$	0,042	RNASE1, CHGA, CNGB1, GNAI2, GNAT1, HSPA8, MYOD1, PIK3R1, CSNK2A2, PNP, POLR2D, ANG
	D018698	Ácido glutâmico	$8,37 \times 10^{-5}$	0,046	CNGB1, GNAT1, HSPA8, ADAMTS4, PNP, ANG
	D008970	Peso molecular	$1,02 \times 10^{-8}$	$2,55 \times 10^{-6}$	ADM, PTH, RNASE1, CHGA, MSTN, GNAT1, HSPA8, MMP13, PIK3R1, SPON1, ADAMTS4, GOT2, MYH9, LYVE1, CALCB, DDR2
	D004789	Ativação Enzimática	$6,72 \times 10^{-7}$	0,0002	PROC, HSPA8, PIK3R1, NDUFS2, COPB1, COPA
Fenômenos e processos	D013379	Especificidade do substrato	$7,71 \times 10^{-6}$	0,002	MB, PTH, RNASE1, CHGA, GPX1, GNAT1, HYAL2, RNASE6, RHOA, TNNC1, APEH, GLRX, ANG
	D010766	Fosforilação	$5,56 \times 10^{-5}$	0,014	MMP3, ADAM2, ADAM9, MMP13, ADAMTS4
	D015394	Estrutura molecular	$6,07 \times 10^{-5}$	0,0152	PTH, GPX1, STAR, CEBPA, HSPA8, MYOD1, PIK3R1, TLR9, USF1, NUCB2, ATF6, SWAP70, ZNF219, POLR2D, YAP1
	D013329	Relação Estrutura/Atividade	0,00014	0,0356	MMP1, MMP3, MMP13, MMP15, MMP7, ADAMTS4

D017433	Estrutura Proteica Secundária	0,00019	0,048	<i>RNASE1, CAST, GPR68, MMP20, RNASE6, GOT2, PNP, TNNC1, APEH, SH3BP4, ATP1A4, MANF, ANG</i>
---------	----------------------------------	---------	-------	--

---

**Tabela Suplementar S6:** Genes candidatos encontrados via recurso de anotação e análise de enriquecimento MeSH para haplótipos da raça Girolando.

ID Gene <sup>1</sup>	ID Gene Ensembl <sup>2</sup>	Haplótipo <sup>3</sup>	Região (pb) <sup>4</sup>		Descrição <sup>2</sup>	Classificação <sup>2</sup>
			Base inicial	Base final		
<i>POLR2D</i>	ENSBTAG00000053244	GR2H1	4.623.450	4.644.243	Subunidade D da RNA polimerase II	
<i>PROC</i>	ENSBTAG00000003950	GR2H1	5.102.056	5.112.985	Proteína C, inativador dos fatores de coagulação Va e VIIIa	
<i>MSTN</i>	ENSBTAG00000011808	GR2H1	6.278.864	6.285.486	Miostatina	
<i>DDR2</i>	ENSBTAG00000004885	GR3H3	6.671.274	6.853.263	Receptor de domínio discoidina tirosina quinase 2	
<i>ATF6</i>	ENSBTAG00000005227	GR3H3	7.616.413	7.846.328	Ativando o fator de transcrição 6	
<i>MPZ</i>	ENSBTAG00000033835	GR3H3	8.207.973	8.216.078	Proteína mielina zero	
<i>NDUFS2</i>	ENSBTAG00000002203	GR3H3	8.290.123	8.304.898	NADH: subunidade central S2 da ubiquinona oxidoreductase	
<i>ADAMTS4</i>	ENSBTAG00000013210	GR3H3	8.303.802	8.311.860	Metalopeptidase ADAM com motivo 4 de trombospondina tipo 1	
<i>USF1</i>	ENSBTAG00000017873	GR3H3	8.436.890	8.442.549	Fator de transcrição upstream 1	
<i>F11R</i>	ENSBTAG00000017846	GR3H3	8.464.412	8.488.977	Receptor F11	Gene codificador de proteína
<i>ITLN2</i>	ENSBTAG00000052308	GR3H4	8.628.057	8.634.971	Intelectina 2	
<i>COPA</i>	ENSBTAG00000004333	GR3H4	9.350.633	9.394.688	Subunidade alfa do complexo de revestimento COPI	
<i>ATP1A4</i>	ENSBTAG00000038523	GR3H4	9.487.900	9.520.439	Subunidade alfa 4 de transporte de ATPase Na <sup>+</sup> /K <sup>+</sup>	
<i>ATP1A2</i>	ENSBTAG00000010551	GR3H4	9.528.257	9.552.859	Subunidade alfa 2 de transporte de ATPase Na <sup>+</sup> /K <sup>+</sup>	
<i>SAG</i>	ENSBTAG00000021480	GR3H5	113.023.696	113.054.214	Parada visual do antígeno S	
<i>SPP2</i>	ENSBTAG00000002030	GR3H5	113.703.201	113.730.899	Fosfoproteína 2 secretada	
<i>SH3BP4</i>	ENSBTAG00000013869	GR3H5	114.604.067	114.633.955	Proteína 4 de ligação ao domínio SH3	
<i>HMOX1</i>	ENSBTAG00000015582	GR5H6	73.623.453	73.630.534	Heme oxigenase 1	
<i>MB</i>	ENSBTAG00000005333	GR5H6	73.814.165	73.824.954	Mioglobina	
<i>MYH9</i>	ENSBTAG00000010402	GR5H6	74.719.956	74.805.905	Cadeia pesada 9 da miosina	
<i>RAB28</i>	ENSBTAG00000003218	GR6H7	108.561.459	108.697.586	RAB28, membro da família RAS onco	
<i>GLRX</i>	ENSBTAG00000038186	GR7H8	-	-	-	

<i>PCSK1</i>	ENSBTAG00000020843	GR7H8	95.745.639	95.790.112	Pró-proteína convertase subtilisina/kexin tipo 1	
<i>CAST</i>	ENSBTAG00000000874	GR7H8	95.956.377	96.167.151	Calpastatina	
<i>LNPEP</i>	ENSBTAG00000019900	GR7H9	96.372.198	96.467.677	Leucil e cistinil aminopeptidase	
<i>RPGRIP1</i>	ENSBTAG00000004694	GR10H11	-	-	-	
<i>ZNF219</i>	ENSBTAG00000021008	GR10H11	25.986.954	26.001.853	Proteína de dedo de zinco 219	
<i>RNASE13</i>	ENSBTAG00000046189	GR10H11	26.063.088	26.071.809	Ribonuclease A membro da família 13 (inativo)	
<i>RNASE1</i>	ENSBTAG00000025663	GR10H11	26.223.844	26.225.388	Ribonuclease, família RNase A, 1 (pancreática)	
<i>BRB</i>	ENSBTAG00000004950	GR10H11	26.256.214	26.257.979	Ribonuclease cerebral	
<i>RNASE6</i>	ENSBTAG00000008792	GR10H11	26.369.715	26.380.695	Ribonuclease Um membro da família k6	
<i>RNASE4</i>	ENSBTAG00000019612	GR10H11	26.391.103	26.392.497	Ribonuclease Um membro da família 4	
<i>ANG</i>	ENSBTAG00000019612	GR10H11	26.396.434	26.397.000	Angiogenina, ribonuclease, família RNase A, 5	
<i>ANG2</i>	ENSBTAG00000046668	GR10H11	26.401.251	26.413.235	Angiogenina 2	
<i>RNASE10</i>	ENSBTAG00000032869	GR10H11	26.581.549	26.606.014	Ribonuclease 1 membro da família 10 (inativo)	
<i>PNP</i>	ENSBTAG00000012317	GR10H11	26.632.023	26.650.726	Purina nucleosídeo fosforilase	Gene
<i>MMP13</i>	ENSBTAG00000015059	GR15H12	5.670.871	5.681.601	Metalopeptidase de matriz 13	codificador de
<i>MMP3</i>	ENSBTAG00000037768	GR15H12	5.779.177	5.785.763	Metalopeptidase de matriz 3	proteína
<i>MMP1</i>	ENSBTAG00000048029	GR15H12	5.811.627	5.820.952	Metalopeptidase de matriz 1 (colagenase intersticial)	
<i>MMP8</i>	ENSBTAG00000046732	GR15H12	5.933.094	5.944.195	Metalopeptidase de matriz 8	
<i>MMP20</i>	ENSBTAG00000014251	GR15H12	6.031.818	6.089.843	Metalopeptidase de matriz 20	
<i>MMP7</i>	ENSBTAG00000004594	GR15H12	6.143.363	6.153.606	Metalopeptidase de matriz 7	
<i>YAP1</i>	ENSBTAG00000039307	GR15H13	6.501.783	6.634.647	Regulador transcricional associado YAP1	
<i>TRPC6</i>	ENSBTAG00000011227	GR15H13	7.313.317	7.473.046	Potencial receptor transitório canal catiônico subfamília C membro 6	
<i>PGR</i>	ENSBTAG00000024648	GR15H13	7.854.787	7.966.985	Receptor de progesterona	
<i>AMOTL1</i>	ENSBTAG00000018945	GR15H16	15.542.579	15.735.942	Angiomotina like 1	
<i>HSPA8</i>	ENSBTAG00000013162	GR15H18	-	-	-	
<i>MYOD1</i>	ENSBTAG00000002216	GR15H18	34.794.122	34.796.767	Diferenciação miogênica 1	
<i>KCNJ11</i>	ENSBTAG00000016170	GR15H18	35.113.703	35.116.350	Canal de retificação interna de potássio subfamília J membro 11	

<i>NUCB2</i>	ENSBTAG00000017468	GR15H18	35.156.671	35.179.314	Nucleobindina 2	
<i>CALCB</i>	ENSBTAG00000001516	GR15H19	37.806.147	37.809.438	Polipeptídeo beta relacionado à calcitonina	
<i>COPB1</i>	ENSBTAG00000006556	GR15H19	38.233.156	38.267.714	Subunidade beta 1 do complexo de revestimento COPI	
<i>RRAS2</i>	ENSBTAG00000011905	GR15H19	38.367.194	38.460.551	Relacionado a RAS 2	
<i>SPON1</i>	ENSBTAG00000009150	GR15H19	38.457.984	38.773.001	Espondina 1	
<i>PTH</i>	ENSBTAG00000019080	GR15H19	39.190.651	39.193.187	Hormônio da paratireóide	
<i>EIF4G2</i>	ENSBTAG00000020308	GR15H20	41.871.084	41.881.439	Fator de iniciação da tradução eucariótica 4 gama 2	
<i>IRAG1</i>	ENSBTAG00000007129	GR15H20	41.972.410	42.098.627	Receptor de inositol 1,4,5-trifosfato associado 1	
<i>LYVE1</i>	ENSBTAG00000000802	GR15H20	42.102.179	42.116.460	Receptor 1 de hialuronano endotelial de vaso linfático	
<i>ADM</i>	ENSBTAG00000021048	GR15H20	42.377.322	42.382.803	Adrenomedulina	
<i>SWAP70</i>	ENSBTAG00000001987	GR15H20	42.918.200	43.060.368	Comutação da subunidade do complexo de células B SWAP70	
<i>SLC12A3</i>	ENSBTAG00000006630	GR18H22	-	-	-	
<i>CNGB1</i>	ENSBTAG00000005702	GR18H22	25.813.506	25.880.192	Subunidade beta 1 do canal controlado por nucleotídeo cíclico	Gene codificador de proteína
<i>MMP15</i>	ENSBTAG00000012919	GR18H22	25.939.947	25.960.720	Metalopeptidase de matriz 15	
<i>CSNK2A2</i>	ENSBTAG00000013784	GR18H22	26.016.116	26.051.022	Caseína quinase 2 alfa 2	
<i>GOT2</i>	ENSBTAG00000007172	GR18H22	26.447.740	26.471.481	Transaminase glutâmico-oxaloacética 2	
<i>RGS9BP</i>	ENSBTAG00000039441	GR18H23	43.154.652	43.155.365	Regulador da proteína de ligação à sinalização da proteína G9	
<i>PIK3R1</i>	ENSBTAG00000010989	GR20H24	11.398.588	11.408.015	Subunidade reguladora 1 da fosfoinositídeo-3-quinase	
<i>NLN</i>	ENSBTAG00000016903	GR20H24	13.752.971	13.853.634	Neurolisina	
<i>GPR68</i>	ENSBTAG00000006738	GR21H27	55.998.072	56.036.571	Receptor 68 acoplado à proteína G	
<i>FBLN5</i>	ENSBTAG00000018123	GR21H27	56.567.085	56.659.651	Fibulina 5	
<i>ATXN3</i>	ENSBTAG00000005786	GR21H27	56.760.080	56.794.358	Ataxina 3	
<i>NDUFB1</i>	ENSBTAG00000015156	GR21H27	56.801.799	56.809.044	NADH: subunidade B1 da ubiquinona oxidoreductase	
<i>LGMN</i>	ENSBTAG00000046979	GR21H27	57.379.065	57.414.855	Legumain	
<i>CHGA</i>	ENSBTAG00000009836	GR21H27	57.571.750	57.585.040	Cromogranina A	

<i>CACNA1D</i>	ENSBTAG00000010026	GR22H28	47.162.857	47.658.557	Subunidade alfa1 D do canal dependente de voltagem de cálcio	
<i>PRKCD</i>	ENSBTAG00000008719	GR22H28	47.750.240	47.780.548	Proteína quinase C delta	
<i>TNNC1</i>	ENSBTAG000000045757	GR22H28	48.431.431	48.437.749	Troponina C1, tipo esquelético e cardíaco lento	
<i>TLR9</i>	ENSBTAG00000018198	GR22H28	48.676.670	48.680.933	Pedágio como receptor 9	
<i>MANF</i>	ENSBTAG000000031797	GR22H28	-	-	-	
<i>HYAL2</i>	ENSBTAG00000000484	GR22H28	50.021.694	50.029.994	Hialuronidase 2	
<i>GNAI2</i>	ENSBTAG00000020645	GR22H28	50.099.986	50.120.522	Subunidade alfa I2 da proteína G	
<i>GNAT1</i>	ENSBTAG00000018020	GR22H28	50.152.601	50.161.258	Subunidade alfa transducina 1 da proteína G	Gene codificador de proteína
<i>APEH</i>	ENSBTAG00000011583	GR22H28	50.512.079	50.524.496	Hidrolase de acilaminoacil-peptídeo	
<i>RHOA</i>	ENSBTAG000000004279	GR22H28	50.701.226	50.751.132	Membro da família homóloga ras A	
<i>GPX1</i>	ENSBTAG000000054195	GR22H28	50.752.554	50.753.718	Glutaciona peroxidase 1	
<i>ERLIN2</i>	ENSBTAG00000017831	GR27H29	33.016.693	33.033.829	Jangada lipídica ER associada 2	
<i>STAR</i>	ENSBTAG00000033345	GR27H29	33.312.972	33.321.495	Proteína reguladora aguda esteroideogênica	
<i>ADAM9</i>	ENSBTAG00000017455	GR27H29	-	-	-	
<i>ADAM2</i>	ENSBTAG00000009408	GR27H29	34.593.321	34.693.833	Domínio 2 da metalopeptidase ADAM	

<sup>1</sup> ID dos genes candidatos encontrados via MeSH;

<sup>2</sup> ID, descrição e classificação encontrados via Ensembl Genes 103 (DURINCK et al., 2009), usando o *btaurus\_gene\_ensembl*;

<sup>3</sup> Código de nomeação do haplótipo que inclui a raça (GI para Gir e GR para Girolando), o cromossomo e o número de haplótipos dentro da raça (H para haplótipo);

<sup>4</sup> Genoma de referência *Bos taurus* ARS-UCD 1.2 (ROSEN et al., 2020).