UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CAMPUS JABOTICABAL

GENOME SCAN FOR GENOMIC REGIONS, LETHAL HAPLOTYPES, AND TRANSMISSION RATIO DISTORTIONS ASSOCIATED WITH REPRODUCTIVE TRAITS AND PREWEANING MORTALITY IN NELLORE CATTLE

Gustavo Roberto Dias Rodrigues
Animal Scientist

2025

DISS. / RODRIGUES G.R.D. 2025

UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CAMPUS JABOTICABAL

GENOME SCAN FOR GENOMIC REGIONS, LETHAL HAPLOTYPES, AND TRANSMISSION RATIO DISTORTIONS ASSOCIATED WITH REPRODUCTIVE TRAITS AND PREWEANING MORTALITY IN NELLORE CATTLE

Student: Gustavo Roberto Dias Rodrigues
Advisor: Prof. Dr. Maria Eugênia Zerlotti Mercadante
Coadvisor: Dr. Joslaine Noely dos Santos Gonçalves Cyrillo
Coadvisor: Dr. Lúcio Flávio Macedo Mota

Master's thesis presented to the Graduate program in Animal Science at the School of Agricultural and Veterinary Science – UNESP, Jaboticabal campus, as part of the requirements for obtaining the title of Master's in animal science.

2025

R696g

Rodrigues, Gustavo Roberto Dias

Genome scan for genomic regions, lethal haplotypes, and transmission ratio distortions associated with reproductive traits and preweaning mortality in Nellore cattle / Gustavo Roberto Dias Rodrigues. -- Jaboticabal, 2025

168 p.

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

Orientador: Maria Eugênia Zerlotti Mercadante

Coorientador: Joslaine Noely dos Santos Gonçalves Cyrillo

1. Animal Breeding and Genetics. 2. Genomics. 3. Reproduction. 4. Lethal Alleles. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Dados fornecidos pelo autor(a).

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: GENOME SCAN FOR GENOMIC REGIONS, LETHAL HAPLOTYPES, AND TRANSMISSION RATIO DISTORTIONS ASSOCIATED WITH REPRODUCTIVE TRAITS AND PREWEANING MORTALITY IN NELLORE CATTLE

AUTOR: GUSTAVO ROBERTO DIAS RODRIGUES
ORIENTADORA: MARIA EUGÊNIA ZERLOTTI MERCADANTE
COORIENTADOR: LÚCIO FLÁVIO MACÊDO MOTA
COORIENTADORA: JOSLAINE NOELY DOS SANTOS GONÇALVES CYRILLO
ORIENTADOR DE MOBILIDADE: LUIZ FERNANDO BRITO (Internacional)

Aprovado como parte das exigências para obtenção do Título de Mestre em Ciência Animal, área: Genética e Melhoramento Animal pela Comissão Examinadora:

gov.br Documento assinado digitalmente
MARIA EUGENIA ZERLOTTI MERCADANTE
Data: 09/11/2025 15:23:57-0300
Verifique em https://validar.iti.gov.br

Pesquisadora Dra. MARIA EUGÊNIA ZERLOTTI MERCADANTE (Participaçao Virtual)
IZ - Insituto de Zootecnia / Sertãozinho/SP

Dr. VICTOR BRENO PEDROSA (Participaçao Virtual)
Neogen Corporation / Lincoln/NE/EUA

gov.br Documento assinado digitalmente
FERNANDO SEBASTIAN BALDI REY
Data: 09/11/2025 16:35:58-0300
Verifique em https://validar.iti.gov.br

Prof.Ass.Dr. FERNANDO SEBASTIAN BALDI REY (Participaçao Virtual)
Departamento de Medicina Veterinária / FZEA USP Pirassununga/SP

Jaboticabal, 07 de novembro de 2025.

POTENTIAL IMPACT OF THIS RESEARCH

First study: This study advances the understanding of the genomic background underlying key reproductive traits in Nellore cattle: conception success, pregnancy loss, stillbirth, and pre-weaning mortality. By identifying genomic regions, candidate genes, and biological pathways associated with these complex traits, the research provides a scientific basis for enhancing genomic selection strategies aimed at improving reproductive efficiency. The results have the potential to inform breeding programs, enabling more accurate selection decisions for traits that are difficult to measure and typically exhibit low heritability. Additionally, the findings serve as reference for future functional studies in Zebu cattle.

Second study: This study provides contributions to the detection of lethal alleles in Nellore cattle by identifying 45 genomic regions harboring putative lethal haplotypes. The annotation of candidate genes and biological pathways associated with sexual development, immune function, energy homeostasis, and embryonic viability underscores the potential role of these genomic regions in causing embryonic losses and compromised reproductive performance. The results highlight detrimental effects of matings between carriers of lethal haplotypes on key reproductive outcomes, including reductions in pregnancy success and increases in pregnancy loss, stillbirths, and pre-weaning mortality. The findings offer a valuable background for implementing genomic strategies to identify and manage carriers, thereby enabling more informed mating decisions to reduce the propagation of lethal alleles in breeding populations.

Third study: This study represents the first comprehensive characterization of genomic regions exhibiting transmission ratio distortion in Nellore cattle, providing novel perceptions into non-Mendelian inheritance mechanisms influencing reproductive outcomes in this breed. By identifying 1,249 genomic regions displaying TRD patterns and highlighting 132 regions as potential carriers of semi-lethal or lethal alleles, the research reveals candidate genes and biological pathways implicated in embryonic development, morphogenesis, and growth regulation. The integration of TRD analyses into breeding programs holds significant potential to enhance reproductive efficiency, mitigate the propagation of deleterious alleles, and improve the sustainability of beef production systems.

AUTHOR'S CURRICULAR DATA

Gustavo Roberto Dias Rodrigues was born in Araguari, Minas Gerais, Brazil, on June 26, 2001. In August 2018, he began his undergraduate studies in Animal Science at the Federal University of Uberlândia (UFU). During his studies, he was awarded undergraduate research scholarships from both the Research Support Foundation of the State of Minas Gerais (FAPEMIG) and the National Council for Scientific and Technological Development (CNPq). He held a leadership role at the Laboratory for Agribusiness Studies (LEA/UFU) and gained practical experience through internships at both the University of São Paulo (USP) and the Beef Cattle Research Center – Institute of Animal Science (IZ) under the supervision of Dr. Joslaine Cyrillo. He obtained his bachelor's degree in Animal Science in June 2023. In August 2023, he started his Master's program in Animal Science at São Paulo State University (UNESP), funded by a scholarship from the São Paulo Research Foundation (FAPESP, grant #2023/11176-4), under the guidance of Dr. Maria E. Z. Mercadante. He also completed a seven-month research internship abroad, supported by FAPESP (grant #2024/01909-7), at Purdue University (West Lafayette, IN, USA), under the supervision of Dr. Luiz F. Brito. During his Master's studies, he actively contributed to multiple research projects and authored articles published in high-impact journals. He was recognized with the "Future Scientist" award in 2023 at the 17th CIIC, organized by Embrapa, IAC, ITAL, IZ, IB, and IP, and received an honorable mention for best abstract at the XI Technical-Scientific Colloquium on One Health, Agricultural Sciences, and the Environment. Additionally, in 2024 he co-authored a study that was honored with the "Fabyano Fonseca e Silva" award for best abstract at the III Symposium on Genetics and Animal Breeding (SIMMelhor). In July 2025, he was nominated for the SBZ Excellence in Graduate Studies Award, which recognizes the top graduate students in Animal Science in Brazil.

‘‘The journey of a thousand miles begins with a single step”

- Lao Tzu

“Remember to look up at the stars and not down at your feet”

- Stephen Hawking

“What would life be if we had no courage to attempt anything?”

- Vicent Van Gogh

I dedicate this achievement to my grandmother, Izamita Naves da Costa, whose love and guidance continue to inspire me.

ACKNOWLEDGEMENTS

I would first like to express my deepest gratitude to my family for their unconditional love and support throughout my life. None of this would have been possible without you. Thank you for your patience, encouragement, and for all the moments that kept me moving forward.

I am sincerely grateful to my advisor, Maria Eugênia, and my co-advisor, Joslaine, who have inspired me and believed in me since 2022, even before I began my master's studies. Your guidance has been fundamental to my personal and professional growth. Words cannot fully express how important you have been in shaping who I am today.

My heartfelt thanks go to my partner, João, who is always by my side and whose support never fails. You are one of the most intelligent people I have ever met, and sharing life with you makes me profoundly happy. Thank you for everything.

I am also immensely grateful to my dear friends Eduarda, Milena, Júlia, Marco Túlio, Ana Carolina, Arícia, Maria Rita, Cassiane, Henrique, and Gustavo Henrique for walking with me through different stages of my master's journey. I appreciate every conversation, every laugh, and every piece of support you gave me, especially during difficult moments (and there were many!).

To all my friends and colleagues from UFU, IZ, UNESP, and Purdue University, thank you for the valuable moments we shared. Each of you has taught me something meaningful that I will carry with me throughout my life, and I hope I was also able to make a positive contribution to your journeys. A special thanks goes to Patrícia and Lúcio, who helped me with some analyses presented in this dissertation.

I would also like to acknowledge my supervisor abroad, Luiz Brito, for his important role during the final stage of my master's. I am grateful for all the opportunities, guidance, advice, and for supervising the fourth chapter presented in this dissertation.

À Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, pelo auxílio financeiro para desenvolvimento desta pesquisa – processos 2023/11176-4 e 2024/01909-7.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

TABLE OF CONTENTS

GENOME SCAN FOR GENOMIC REGIONS, LETHAL HAPLOTYPES, AND TRANSMISSION RATIO DISTORTIONS ASSOCIATED WITH REPRODUCTIVE TRAITS AND PRE-WEANING MORTALITY IN NELLORE CATTLE

ABSTRACT – Reproductive efficiency is one of the main challenges for beef cattle production, since reproductive traits and pre-weaning calf mortality, generally exhibit low heritability, late expression, and, in some cases, are measured only in one sex, thereby limiting genetic progress. In this context, the application of genomic tools represents a promising strategy to elucidate the genetic architecture underlying these traits while simultaneously mitigating the risks associated with increased inbreeding and the dissemination of deleterious variants. Thus, the present study aimed to investigate the genomic basis of reproductive traits and pre-weaning mortality in Nellore cattle through the integration of three approaches: (i) genome-wide association studies (GWAS) with functional annotation, to identify genomic regions, candidate genes, and biological pathways related to complex reproductive phenotypes; (ii) detection of putative lethal haplotypes in the homozygous state and assessment of their effects on the reproductive performance of heterozygous carriers; and (iii) characterization of genomic regions with transmission ratio distortion (TRD), as potential carriers of lethal or semi-lethal alleles. The first study identified genomic regions associated with conception success, pregnancy loss, stillbirth, and pre-weaning mortality, highlighting candidate genes involved in reproduction, immunity, and metabolism. The second study detected 45 genomic regions with a complete absence of homozygous recessive haplotype carriers, suggesting the presence of lethal alleles influencing reproductive processes such as embryonic loss. The third study mapped 1,249 genomic regions with TRD signals, characterized by deviations from Mendelian inheritance at allelic or genotypic levels, of which 132 were considered potential carriers of lethal or semi-lethal alleles based on the high number of underrepresented progeny. Taken together, these findings provide valuable knowledge that could be used for the improvement of genomic selection strategies aimed at mitigating reproductive losses associated with deleterious variants, thereby contributing to reducing their propagation in Nellore cattle.

Keywords: homozygous haplotypes, Mendelian inheritance, pregnancy loss, stillbirth

VARREDURA DO GENOMA PARA REGIÕES GENÔMICAS, HAPLÓTIPOS LETAIS E DISTORÇÕES NA RAZÃO DE TRANSMISSÃO ASSOCIADAS A CARACTERÍSTICAS REPRODUTIVAS E MORTALIDADE PRÉ-DESMAME EM BOVINOS NELORE

RESUMO – A eficiência reprodutiva constitui um dos principais desafios da pecuária de corte, uma vez que características reprodutivas e a mortalidade de bezerros no período pré-desmame, apresentam, em geral, baixa herdabilidade, expressão tardia e, em alguns casos, são mensuradas apenas em um dos sexos, o que limita o progresso genético. Nesse contexto, a aplicação de ferramentas genômicas representa uma estratégia promissora para elucidar a arquitetura genética subjacente a essas características e, simultaneamente, mitigar os riscos associados ao aumento da endogamia e à disseminação de variantes deletérias. Assim, o presente trabalho teve como objetivo geral investigar a base genômica de características reprodutivas e de mortalidade pré-desmame em bovinos Nelore, por meio da integração de três abordagens: (i) estudos de associação genômica ampla (GWAS) com anotação funcional, visando identificar regiões genômicas, genes candidatos e vias biológicas relacionadas a fenótipos reprodutivos complexos; (ii) detecção de haplótipos supostamente letais em estado homozigoto e avaliação de seus efeitos sobre o desempenho reprodutivo de portadores heterozigotos; e (iii) caracterização de regiões com distorção na razão de transmissão (TRD), potenciais portadoras de alelos letais ou semi-letais. Os resultados do primeiro estudo evidenciaram regiões genômicas associadas ao sucesso de concepção, à perda gestacional, à natimortalidade e à mortalidade pré-desmame, destacando genes candidatos envolvidos em processos de reprodução, imunidade e metabolismo. O segundo estudo identificou 45 regiões genômicas com ausência de indivíduos portadores de haplótipos homozigotos recessivos, sugerindo a presença de alelos letais influenciando processos associados a reprodução, como a perda embrionária. O terceiro estudo mapeou 1.249 regiões genômicas com sinais de TRD, caracterizadas por desvios de herança mendeliana alélica ou genotípica, dentre as quais 132 foram consideradas potenciais portadoras de alelos letais ou semi-letais, com base no elevado número de progênies sub-representadas. De forma integrada, esses achados fornecem informações que podem ser utilizadas para o aprimoramento de estratégias

de seleção genômica voltadas à mitigação de perdas reprodutivas associadas a variantes deletérias, contribuindo para reduzir sua propagação em bovinos Nelore.

Palavras-chave: haplótipos homozigotos, herança mendeliana, natimorto, perda gestacional

CHAPTER 1: GENERAL CONSIDERATIONS

1.1. Introduction

The genetic improvement of reproductive performance in beef cattle is challenging due to the inherent complexity of fertility-related traits, which typically exhibit low heritability, are expressed later in life, and often sex-specific, factors that can reduce the accuracy of estimated breeding values and slow genetic progress (Falconer & Mackay, 1996; Toghiani et al., 2017; Olasege et al., 2021). Advances in genomic technologies have transformed breeding strategies by enabling the use of dense panels of single nucleotide polymorphisms (SNPs) distributed across the genome to estimate genomic breeding values with higher accuracy (Meuwissen et al., 2001; VanRaden, 2008). This approach enhances the precision of selection, facilitates the earlier identification of animals with superior genetic potential, and accelerates the rate of genetic gain (Meuwissen et al., 2001; Schaeffer, 2006; Hayes et al., 2009), offering pathways for addressing some of the limitations in improving reproductive performance in cattle.

However, the widespread use of genomic selection has also introduced potential negative impacts that must be considered (Misztal & Lourenco, 2024). One concern is the potential for increased rates of inbreeding due to intensive use of a few top-ranking sires within breeding programs (Makanjuola et al., 2020; Lozada-Soto et al., 2021). Elevated inbreeding raises the likelihood of homozygosity for deleterious recessive alleles, thereby increasing inbreeding depression in economically important traits and the risk of genetic defects (Howard et al., 2017; Bem et al., 2024; Mulim et al., 2025). Moreover, strong selection for a specific trait can inadvertently promote the spread of deleterious alleles if not carefully managed, particularly when an advantageous variant is in high linkage disequilibrium (LD) with a deleterious allele (VanRaden et al., 2011; Liu et al., 2014; Jenko et al., 2019). In such cases, the beneficial allele and the harmful variant may rise in frequency together because they tend to be inherited as a block if recombination does not occur between them (Hartfield & Otto, 2011; Jenko et al., 2019). As a result, alleles with negative effects on reproductive performance can persist or even increase in prevalence within the population despite their undesirable outcomes.

To address these challenges and mitigate potential risks, several efforts are being made to identify deleterious genetic variants associated with reproductive performance in beef cattle. Genome-wide association studies (GWAS) have become a fundamental tool to detect statistical associations between SNP markers and reproductive phenotypes, uncovering genomic regions and potential candidate genes influencing traits of interest (Sbardella et al., 2021; Rodrigues et al., 2024; Ogunbawo et al., 2025). Another approach involves the detection of putative lethal haplotypes through the absence of homozygous carriers for recessive haplotypes (VanRaden et al., 2011). This method has proven effective in identifying recessive lethal alleles that may cause embryonic and pregnancy losses, leading to practical applications in breeding decisions to avoid carrier-to-carrier matings (Schmidt et al., 2023; Rodrigues et al., 2025a). Additionally, transmission ratio distortion (TRD) analyses have emerged as a novel approach in livestock species for identifying potential candidates for semi-lethal or lethal alleles, revealing genomic regions where observed inheritance patterns deviate from expected Mendelian ratios and thereby indicating possible deleterious effects on fertility or embryonic viability (Id-Lahoucine et al., 2023a; Rodrigues et al., 2025b).

Despite these advances, some knowledge gaps remain in Nellore regarding the genomic factors influencing reproductive performance. Few studies focused on identifying genomic regions associated with conception success, pregnancy loss, stillbirth, and pre-weaning calf mortality through GWAS in Nellore cattle. Furthermore, although Schmidt et al. (2023) reported 30 candidates for lethal haplotypes in Nellore cattle, this finding still requires validation across independent populations to confirm their prevalence in other herds. Moreover, to the best of our knowledge, prior to this study, no comprehensive analyses of transmission ratio distortion (TRD) had been conducted in Nellore cattle, leaving an important gap in understanding potential allelic and genotypic deviations from Mendelian inheritance that might signal the presence of deleterious alleles affecting fertility and embryonic viability.

Hence, the objectives of this Master's thesis were to: (i) perform GWAS and functional annotation analyses to identify genomic regions, candidate genes, biological processes, and metabolic pathways associated with conception success, pregnancy loss, stillbirth, and pre-weaning calf mortality in Nellore cattle; (ii) detect putative lethal haplotypes in closed experimental selection lines of Nellore cattle based on the absence of recessive carriers of homozygous haplotypes, and to evaluate their

potential impacts on reproductive performance; and (iii) conduct a comprehensive TRD analysis in Nellore cattle to uncover genomic regions exhibiting non-Mendelian inheritance patterns that may harbor putative semi-lethal or lethal alleles affecting reproduction.

1.2. Literature Review

1.2.1 Genome-wide association studies and functional annotation for reproductive traits in cattle

Understanding the genetic aspects of economically relevant and biologically complex traits has become essential for advancing cattle breeding, and the selection for improved reproductive traits remains among the most challenging goals. Fertility traits, such as conception success, pregnancy loss, calving interval, days to calving, and age at first calving, are typically characterized by low to moderate heritability, late expression, and substantial environmental influence, factors that reduce the effectiveness of selection (Berry et al., 2014; Lucy, 2019; Bessa et al., 2021). These limitations have emphasized the need for genomic approaches capable of revealing the underlying genetic background of such traits, making GWAS one of the most effective tools for this purpose. GWAS provides a framework to dissect the genomic architecture of complex phenotypes by evaluating associations between SNPs and trait variability (Hirschhorn & Daly, 2005). By exploiting patterns of LD between genetic markers and causal mutations, GWAS enables the identification of quantitative trait loci (QTL) without prior assumptions about their genomic location (Johnson et al., 2021; Uffelmann et al., 2021; Merrick et al., 2022). The development of high-density SNP panels and advances in statistical methodologies, such as single-step Genomic Best Linear Unbiased Prediction (ssGBLUP), have greatly enhanced the resolution of GWAS, allowing the detection of genomic regions that explain portions of additive genetic variance (Wang et al., 2012), and enabling the identification of specific SNPs with significant effects on economically important traits (Aguilar et al., 2019).

Importantly, GWAS findings are often complemented by functional analyses of significant loci, including gene annotation and comprehensive enrichment analyses of biological processes, molecular functions, and metabolic pathways (Gallagher & Chen-Plotkin, 2018). Once significant SNPs or QTL regions are identified, they are typically mapped to nearby or overlapping genes, followed by functional annotation to

characterize their potential biological roles (Silberstein et al., 2021). This process often leverages databases such as Ensembl (https://www.ensembl.org/), Nacional Library of Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/), or UniProt (https://www.uniprot.org/) to gather information on gene functions, expression patterns, and known associations with phenotypes. Subsequently, enrichment analyses can be performed to evaluate whether specific biological processes, molecular functions, or cellular components are overrepresented among the candidate genes (Gallagher & Chen-Plotkin, 2018; Silberstein et al., 2021). Commonly used frameworks include the Gene Ontology (GO, https://geneontology.org/), which classifies gene functions hierarchically (e.g., biological process, molecular function, cellular component) (Ashburner et al., 2000), and the Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.kegg.jp/), which contextualizes genes within signaling cascades, metabolic networks, and regulatory pathways (Kanehisa & Goto, 2000). These post-GWAS analyses transform a list of statistically significant SNP markers or genomic windows explaining a substantial proportion of additive genetic variance into a biological context, facilitating the understanding of mechanisms underlying trait variability and prioritization of candidate genes for functional validation (Hou & Zhao, 2013; Cai et al., 2018).

The first applications of GWAS to reproductive traits in cattle were predominantly carried out in *Bos taurus taurus* dairy breeds. Initial investigations focused on fertility phenotypes such as daughter pregnancy rate, conception rate, calving interval, and days open, traits characterized by low heritability and a complex polygenic basis (Pryce et al., 2010; Cole et al., 2011). Cole et al. (2011) performed a GWAS for 31 production, health, reproduction, and body conformation traits in US Holstein cows using a 50K SNP panel, identifying loci with pleiotropic effects across production and fertility. Significant associations were reported between the *INSR* gene on *Bos taurus* autosome (BTA) 7 and *GRIA3* on BTAX with daughter pregnancy rate and productive life, as well as a genomic region on BTA18 (53.9–58.7 Mb) linked to calving ease and stillbirth. Parker-Gaddis et al. (2016) identified loci of moderate effect for fertility traits in Holstein cattle, including daughter pregnancy rate, heifer conception rate, and cow conception rate, with signals overlapping previously reported QTL and mapping to regions near the genes *DAZL* (important role in gametogenesis) and *BAX* (absence can result in infertility) (Parker-Gaddis et al., 2016). Chen et al. (2022) utilized whole-genome sequence (WGS) imputation to perform a GWAS

encompassing 5.6 million SNPs across 18 fertility and reproductive traits measured in heifers, cows, and sires, revealing key genomic regions on BTA22 and BTA23 for heifer traits, BTA8 and BTA17 for cow traits, and BTA4, BTA7, BTA17, BTA22, BTA25, and BTA28 for sire traits. Functional annotation of significant SNPs highlighted immune response and fatty acid metabolism pathways as major biological processes influencing the expression of fertility and reproductive phenotypes in Holstein cattle (Chen et al., 2022).

In addition to studies conducted in dairy breeds, some GWAS have also investigated reproductive traits in beef cattle. Akanno et al. (2015) conducted a GWAS in crossbred beef heifers composed of Angus, Charolais, and Hereford, evaluating pre-breeding weight, pregnancy rate, age at first calving, calving difficulty, and calf growth performance. Significant SNPs were identified across 15 chromosomes (BTA1, BTA5, BTA7, BTA9, BTA13–BTA16, BTA19, BTA21, BTA24, BTA25, and BTA27–BTA29), with pleiotropic loci on BTA1, BTA5, and BTA24 influencing multiple traits. Speidel et al. (2018) focused on Red Angus cattle, identifying nine markers for heifer pregnancy (BTA1, BTA11, BTA13, BTA23, BTA29) and twelve for stayability (BTA6, BTA8, BTA9, BTA12, BTA15, BTA18, BTA22, BTA23). Sweet et al. (2020) investigated male fertility in crossbred beef bulls (Angus, Simmental, Piedmontese, Gelbvieh, Charolais, and Limousin) using the weighted ssGBLUP (WssGBLUP) approach, detecting eight genomic windows on BTA9, BTA10, BTA20, BTA24, and BTA29 explaining more than 1% of the genetic variance for scrotal circumference and five on BTA9, BTA13, BTA20, and BTA24 for sperm motility. Stegemiller et al. (2021) analyzed fertility traits in crossbred heifers sired by Angus, Hereford, Simmental, SimAngus, and Shorthorn, reporting loci on BTA2, BTA3, and BTA23 for antral follicle count and on BTA2, BTA8, BTA10, and BTA11 for reproductive tract score.

Within *Bos taurus indicus*, some GWAS have been conducted in Nellore cattle to dissect the genomic background of reproductive traits under tropical conditions. Costa et al. (2015) applied a Bayesian GWAS to evaluate heifer rebreeding and age at first calving, identifying 42 SNPs associated with heifer rebreeding and 19 with age at first calving, which together explained 11.35% and 6.42% of the phenotypic variance, respectively. Irano et al. (2016) evaluated early pregnancy in heifers and scrotal circumference in bulls through GWAS, identifying genomic windows on BTA5, BTA6, BTA7, BTA14, BTA18, BTA21, and BTA27 associated with early pregnancy, and genomic windows on BTA4, BTA8, BTA11, BTA13, BTA14, BTA19, BTA22, and

BTA23 associated with scrotal circumference, which together explained 7.91% and 6.78% of the genetic variance, respectively. Melo et al. (2017) applied the WssGBLUP approach to heifer rebreeding and number of calvings by 53 months of age, identifying key genomic regions including BTA23 (25–26 Mb), where *MHC* class II genes were associated with pregnancy success, alongside other windows enriched for lipid metabolism and immune response pathways. Dias et al. (2024) performed a GWAS and functional annotation for calving interval in Nellore cattle, and the results highlighted genes related to cell cycle (*RAD21*), oocyte function (*LHX8*, *CLPX*), and endocrine regulation (*LRRFIP2*, *GPR158*).

These studies provide evidence that reproductive traits in cattle populations are highly polygenic and influenced by a wide range of candidate genes, biological processes, and metabolic pathways that still need to be better understood.

1.2.2. Identification of putative lethal haplotypes

Genetic defects with deleterious effects on fertility have long been a critical concern in cattle breeding, as the widespread use of elite sires and the resulting increase in inbreeding can inadvertently elevate the frequency of recessive alleles (VanRaden et al., 2011; Howard et al., 2017). When present in homozygous state, some of these variants can lead to early embryonic loss, stillbirth, congenital malformations, or other reproductive inefficiencies, ultimately compromising genetic progress and herd productivity (Cole et al., 2016; Wu et al., 2019). Historically, the identification of such defects relied on pedigree analyses to trace shared ancestry and on phenotypic screening to recognize evident syndromes, as exemplified by descriptions of complex vertebral malformation (Agerholm et al., 2001; Thomsen et al., 2006) and brachyspina syndrome (Charlier et al., 2012) in Holstein cattle. While these methods enabled the identification of some major defects, they were inherently limited in scope and largely incapable of capturing variants with subtle, early acting, or embryonically lethal effects (VanRaden et al., 2011). Early embryonic loss, in particular, represents one of the most challenging manifestations of recessive lethal alleles, as it typically occurs before pregnancy can be clinically confirmed and do not always show clinical signs (Diskin & Morris, 2008; Fritz et al., 2013). This complexity hinders the identification of causal variants through conventional recording systems

and phenotypic observations, allowing many recessive alleles causing early embryonic lethality to persist undetected for generations, reducing fertility (VanRaden et al., 2011).

The advent of dense SNP genotyping and haplotype-based analyses provided a major breakthrough in detecting recessive lethals by enabling the identification of genomic segments missing in the homozygous state across large populations, an indirect indicator of alleles incompatible with viability. VanRaden et al. (2011) proposed a genome-wide methodology to identify haplotypes exerting deleterious effects on fertility, relying solely on pedigree and genomic information, without the necessity of observing affected embryos or individuals. This approach was based on the hypothesis that, when a haplotype segregates at moderate to high frequency within a population, homozygous individuals would be expected to occur unless homozygosity is incompatible with viability. Accordingly, haplotypes occurring at moderate or high frequencies but with no homozygous individuals can be considered putative carriers of recessive lethal mutations associated with embryonic or perinatal lethality. To test this hypothesis, VanRaden et al. (2011) analyzed genomic datasets from Holstein, Jersey, and Brown Swiss cattle, leading to the identification of three major Holstein haplotypes (HH1, HH2, and HH3), as well as BH1 in Brown Swiss.

Following the approach introduced by VanRaden et al. (2011), some studies have expanded haplotype-based screening for recessive lethal haplotypes in livestock species. Häggman & Uimari (2017) evaluated Yorkshire pigs, identifying 26 putative lethal haplotypes across 12 chromosomes, including a region on chromosome 8 (107.0–113.3 Mb) associated with increased stillbirths. Within this region, three candidate genes (*MAD2L1*, *FGF2* and *ANXA5*) were highlighted, each with important biological roles in embryo implantation and placental function, suggesting their involvement in reproductive failures. Similarly, Howard et al. (2017) examined a Large White pig population and identified six chromosomal regions harboring putative lethal recessive haplotypes that significantly reduced the total number born in carrier-to-carrier matings.

In dairy cattle, Wu et al. (2019) analyzed Nordic Holsteins and identified nine novel homozygous-deficient haplotypes, confirmed previously reported lethal genomic regions such as HH3, and demonstrated through reduced insemination success in carrier-to-carrier matings that early embryonic mortality was the likely cause of homozygote deficiency. Häfliger et al. (2022) used whole-genome sequencing in

Swiss Holsteins to identify four new candidate variants for lethal haplotypes (HH13, HH21, HH25 and HH35), implicating genes involved in embryonic development and immune signaling. Al-Khudhair et al. (2024) discovered a mutation within a common haplotype on BTA16 (78.7–80.7 Mb) associated with Holstein calf muscle weakness, identifying the *CACNA1S* gene as the likely causal mutation. Ask-Gullstrand et al. (2024) evaluated carrier status for ten known recessive defects in Holstein and Nordic Red Dairy Cattle, showing that at-risk matings for HH3 increased pregnancy loss and reduced conception rates, reaffirming HH3 as a major cause of early embryonic loss.

In beef cattle studies, Hoff et al. (2017) performed a comprehensive analysis in Angus cattle to identify homozygous-deficient haplotypes, reporting seven novel putative lethal haplotypes across the genome. Jenko et al. (2019) examined Aberdeen Angus, Charolais, Hereford, Limousin and Simmental cattle, detecting one lethal haplotype on BTA16 in Simmental and two semi-lethal haplotypes on BTA14 in Aberdeen Angus, with population frequencies ranging from 8.8% to 15.2%. These haplotypes were also associated with reduced conception rates and exhibited pleiotropic effects on production traits, exemplifying how lethal alleles can persist in populations through heterozygote advantage.

The identification of candidate lethal haplotypes in *Bos taurus indicus*, specifically for Nellore cattle, has advanced in recent years, with studies revealing putative lethal haplotypes that may compromise reproductive efficiency. Schmidt et al. (2023) conducted a comprehensive genome-wide scan in Nellore cattle, evaluating the absence of homozygous recessive haplotypes in more than 62,000 genotyped animals imputed into a high-density SNP panel (770 K). The authors reported 30 haplotypes segregating at moderate to high frequencies (higher than 2.0% in the population) but with no observed homozygous carriers, suggesting the presence of recessive lethal alleles. Several of these haplotypes were associated with negative effects on reproductive traits such as heifer rebreeding, post-natal mortality, and stayability. Rodrigues et al. (2025a) also searched for potential lethal haplotypes in Nellore cattle by analyzing closed experimental selection lines. Their study identified 45 genomic regions harboring putative lethal haplotypes, overlapping 360 annotated genes, and enriched for pathways linked to ovarian steroidogenesis, oocyte meiosis, insulin signaling, and immune function. Importantly, Rodrigues et al. (2025a) reported that matings between heterozygous carriers of putative lethal haplotypes had

reductions in pregnancy success and increases in pregnancy loss, stillbirth, and pre-weaning mortality.

An important convergence between Schmidt et al. (2023) and Rodrigues et al. (2025a) was the detection of three haplotypes consistently identified in both studies as putative lethal haplotypes for Nellore cattle, despite differences in population sampling. These haplotypes were 998.14 (DNA segment 998, haplotype 14, BTA7, 36.55 to 37.25 Mb), 2213.8 (DNA segment 2213, haplotype 8, BTA17, 65.95 to 66.47 Mb), and 2620.22 (DNA segment 2620, haplotype 22, BTA22, 46.69 to 47.39 Mb). The recurrence of these genomic regions reinforces the evidence that they may harbor variants exerting lethal effects in the homozygous state. In Nellore cattle, they constitute prime targets for fine mapping and functional validation using whole genome sequencing and transcriptomic approaches to clarify their potential association with embryonic lethality (Schmidt et al., 2023; Rodrigues et al., 2025a).

Together, these studies underscore the value of haplotype-based approaches for uncovering recessive lethal alleles in cattle and highlight the need for functional studies to better understand biological aspects of these genomic regions and mitigate their impact on reproductive efficiency and genetic gain.

1.2.3. Transmission ratio distortion

Transmission ratio distortion (TRD) refers to deviations from the Mendelian inheritance expectations of 50:50 allele transmission from heterozygous parents to their offspring (Casellas et al., 2020; Huang et al., 2013). Under Mendel's law of segregation, each allele carried by a heterozygous parent should have an equal probability of being transmitted; however, some biological processes can disrupt this balance, resulting in unequal representation of alleles in the offspring generation (Friocourt et al., 2023; Strome et al., 2024). This phenomenon, first described in model organisms such as *Drosophila* (Gershenson, 1928; Sandler & Novitski, 1957), was initially observed as deviations from expected Mendelian segregation ratios in experimental crosses. Gershenson (1928) reported the "sex ratio" trait in *Drosophila obscura*, where males carrying certain X-chromosomes produced disproportionately more female offspring, providing one of the earliest demonstrations of biased transmission. Later, Sandler & Novitski (1957) formalized the concept of meiotic drive,

describing how specific chromosomal elements or gametic mechanisms could favor their own transmission to the next generation. From a theoretical perspective, TRD signals provides opportunities to better understand the biological processes influencing reproduction, offering perceptions into gamete biology and embryonic development (Huang et al., 2013; Casellas et al., 2020; Id-Lahoucine, et al., 2023b).

TRD can manifest at different stages of reproduction, reflecting a complex association of biological mechanisms. At the gametic stage, TRD may arise from meiotic drive or gametic selection, processes in which specific alleles influence the production, functionality, or survival of gametes, thereby biasing their transmission before fertilization occurs (Sandler & Novitski, 1957). At the zygotic or embryonic stage, TRD is often associated with recessive lethal or semi-lethal alleles because embryos homozygous for such alleles may fail to implant, undergo early resorption, or die during gestation, leading to the underrepresentation of particular genotypes in the population (Diskin & Morris, 2008; VanRaden et al., 2011; Id-Lahoucine, et al., 2023b). In addition to these mechanisms, recombination can be an important source of TRD, since crossing-over events in regions containing structural variants or chromosomal rearrangements can disrupt expected segregation patterns, adding complexity to the interpretation of TRD signals (Larracuente & Presgraves, 2012; Fishman & Mcintosh, 2019).

One of the most documented manifestations of TRD in livestock species is the absence of homozygous haplotypes, first addressed by VanRaden et al. (2011). This approach exploits dense SNP data to identify haplotypes that are common in the heterozygous state but never observed in the homozygous state, signaling the presence of recessive lethal or semi-lethal alleles. These haplotypes leave a clear TRD signal in the genome, characterized by a deficit of certain genotypes among offspring (VanRaden et al., 2011). From a TRD modeling perspective, this phenomenon is typically captured as genotypic TRD, since the distortion arises not from biased transmission of individual alleles, but from the absence of homozygous genotypes, leading to deviations from expected Mendelian segregation ratios (Id-Lahoucine et al., 2019).

Bayesian modeling has emerged as a powerful and flexible framework for detecting TRD across the genome, enabling inference at both SNP and haplotype levels. Id-Lahoucine et al. (2019) emphasized that the statistical power to capture TRD is largely determined by the number of genotyped trios (sire–dam–offspring) and by

the proportion of heterozygous parents at each locus, as only heterozygous individuals provide information for detecting deviations from Mendelian segregation. Within this framework, TRD effects can be modeled under distinct statistical parameterizations. At the allelic model, allelic TRD quantifies whether one allele from a heterozygous parent (e.g., A or B) is transmitted more frequently than expected (50:50). The allelic TRD model allows estimation of an overall TRD effect (parent-unspecific) that captures any distortion regardless of parental origin, and sire-TRD and dam-TRD effects, which partition paternal and maternal contributions, respectively (Casellas et al., 2014, 2020). In contrast, genotypic TRD examines deviations in the expected frequency of offspring genotypes (AA, AB, BB), reflecting processes such as zygotic or embryonic loss of particular genotypes (e.g., lethal alleles) (Id-Lahoucine et al., 2019). The genotypic TRD model further distinguishes between an additive TRD effect, representing directional deviations in allele transmission, and a dominance TRD effect, which captures distortions arising from non-additive genotype interactions (Casellas et al., 2012, 2020).

Beyond single-marker inference, haplotype-based analyses substantially improve the discovery of TRD regions, particularly for rare or low-frequency causal variants that may be weakly tagged by individual SNP. Id-Lahoucine et al. (2019) applied sliding windows of increasing length (2–20 SNP) to scan the Holstein cattle genome for TRD, demonstrating that longer haplotypes increase detection power and pinpoint the specific haplotype alleles driving distortion. This approach also mitigates bias introduced by genotyping errors, as haplotype segments provide stronger cumulative evidence than individual SNP. Together, these methodological advances establish a statistical framework for the detection of genetic variants showing TRD, enabling the identification of regions that may harbor deleterious alleles influencing fertility, embryonic viability, and reproductive performance in cattle and other livestock species (Id-Lahoucine et al., 2019).

Abdalla et al. (2020) identified 14 putative lethal haplotypes exhibiting genotypic TRD and 12 putative lethal haplotypes showing allelic TRD in turkeys, including instances of a complete absence of homozygous individuals. Functional enrichment analyses linked these regions revealed critical biological pathways such as mitotic spindle assembly, mismatch repair, and embryonic development, pointing to deleterious alleles likely responsible for embryonic lethality (Abdalla et al., 2020). Similarly, Laseca et al. (2023) conducted a comprehensive TRD screening in horses,

detecting 140 SNPs with allelic TRD and 42 SNPs with genotypic TRD. Among these, 63 markers displayed stallion-specific TRD and 41 showed mare-specific TRD, providing evidence of parent-specific TRD in equines. Functional analyses associated these TRD regions with genes involved in spermatogenesis, oocyte division, and hormonal regulation, demonstrating the potential of TRD analyses to uncover deleterious variants impacting fertility in horses (Laseca et al., 2023).

In *Bos taurus taurus*, Id-Lahoucine et al. (2023a) reported 851 genomic regions with decisive TRD evidence in Angus cattle, including 52 genomic regions showing allelic TRD that accounted for more than 1,000 underrepresented offspring. Notably, some of these TRD regions were associated with reductions of up to 15% in heifer pregnancy rates, providing evidence of the association between TRD events and fertility, and highlighting its utility for detecting harmful alleles that compromise reproductive performance in beef cattle (Id-Lahoucine et al., 2023a). In Holstein cattle, Id-Lahoucine et al. (2023b) identified 604 TRD regions across the genome, many overlapping with genes involved in DNA repair, meiotic regulation, and early embryonic survival. These findings reinforce TRD analysis as a powerful approach to reveal hidden recessive defects and deleterious variants that negatively affect fertility in dairy populations (Id-Lahoucine et al., 2023b).

Rodrigues et al. (2025b) identified TRD signals in *Bos taurus indicus*, mapping 1,249 TRD regions across the Nellore cattle genome, including 802 overall TRD regions, 191 parent-specific TRD regions, and 256 genotypic TRD regions. Within these, 73 allelic TRD regions and 59 genotypic TRD regions were identified as potential carriers of semi-lethal or lethal alleles, with some loci showing complete absence of homozygotes from specific matings and over 1,000 underrepresented offspring. Functional enrichment analyses revealed strong associations with genes involved in embryo development, morphogenesis, and reproductive processes, and highlighted overlaps with previously reported QTL for reproduction and production traits in cattle (Rodrigues et al., 2025b).

Overall, TRD analyses have proven to be a valuable framework for uncovering hidden deleterious variants that compromise fertility and embryonic viability in livestock (Abdalla et al., 2020; Id-Lahoucine et al., 2023a). TRD studies can not only enhance the understanding of non-Mendelian inheritance patterns but also provide knowledge for improving reproductive efficiency and the sustainability of breeding programs by mitigating the spread of deleterious alleles.

References

Abdalla, E. A., Id-Lahoucine, S., Cánovas, A., Casellas, J., Schenkel, F. S., Wood, B. J., & Baes, C. F. (2020). Discovering lethal alleles across the turkey genome using a transmission ratio distortion approach. *Animal Genetics*, *51*(6), 876–889. https://doi.org/10.1111/age.13003

Agerholm, J. S., Bendixen, C., Andersen, O., & Arnbjerg, J. (2001). Complex vertebral malformation in Holstein calves. *Journal of Veterinary Diagnostic Investigation*, *13*(4), 283–289.

Aguilar, I., Legarra, A., Cardoso, F., Masuda, Y., Lourenco, D., & Misztal, I. (2019). Frequentist p-values for large-scale-single step genome-wide association, with an application to birth weight in American Angus cattle. *Genetics Selection Evolution*, *51*(1), 28. https://doi.org/10.1186/s12711-019-0469-3

Akanno, E. C., Plastow, G., Fitzsimmons, C., Miller, S. P., Baron, V., Ominski, K., Basarab, J. A., & Civetta, A. (2015). Genome-wide association for heifer reproduction and calf performance traits in beef cattle. *Genome*, *58*(12), 549–557. https://doi.org/10.1139/gen-2015-0031

Al-Khudhair, A., VanRaden, P. M., Null, D. J., Neupane, M., McClure, M. C., & Dechow, C. D. (2024). New mutation within a common haplotype is associated with calf muscle weakness in Holsteins. *Journal of Dairy Science*, *107*(6), 3768–3779. https://doi.org/10.3168/jds.2023-24121

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene Ontology: tool for the unification of biology The Gene Ontology Consortium*. *Nature Genetics*, *25*(1), 25–29. http://www.flybase.bio.indiana.edu

Ask-Gullstrand, P., Strandberg, E., Båge, R., Rius-Vilarrasa, E., & Berglund, B. (2024). The effect of genetic defects on pregnancy loss in Swedish dairy cattle. *Journal of Dairy Science*, *107*(5), 2999–3005. https://doi.org/10.3168/jds.2023-24159

Bem, R. D., Benfica, L. F., Silva, D. A., Carrara, E. R., Brito, L. F., Mulim, H. A., Borges, M. S., Cyrillo, J. N. S. G., Canesin, R. C., Bonilha, S. F. M., & Mercadante, M. E. Z. (2024). Assessing different metrics of pedigree and genomic

inbreeding and inbreeding effect on growth, fertility, and feed efficiency traits in a closed-herd Nellore cattle population. *BMC Genomics*, *25*(1), 738. https://doi.org/10.1186/s12864-024-10641-3

Berry, D. P., Wall, E., & Pryce, J. E. (2014). Genetics and genomics of reproductive performance in dairy and beef cattle. *Animal*, *8*(1), 105–121. https://doi.org/10.1017/S1751731114000743

Bessa, A. F. de O., Duarte, I. N. H., Rola, L. D., Bernardes, P. A., Gonzaga Neto, S., Lôbo, R. B., Munari, D. P., & Buzanskas, M. E. (2021). Genetic evaluation for reproductive and productive traits in Brahman cattle. *Theriogenology*, *173*, 261–268. https://doi.org/10.1016/j.theriogenology.2021.08.008

Cai, Z., Guldbrandtsen, B., Lund, M. S., & Sahana, G. (2018). Prioritizing candidate genes post-GWAS using multiple sources of data for mastitis resistance in dairy cattle. *BMC Genomics*, *19*(1), 656. https://doi.org/10.1186/s12864-018-5050-x

Casellas, J., Gularte, R. J., Farber, C. R., Varona, L., Mehrabian, M., Schadt, E. E., Lusis, A. J., Attie, A. D., Yandell, B. S., & Medrano, J. F. (2012). Genome scans for transmission ratio distortion regions in mice. *Genetics*, *191*(1), 247–259. https://doi.org/10.1534/genetics.111.135988

Casellas, J., Manunza, A., Mercader, A., Quintanilla, R., & Amills, M. (2014). A flexible bayesian model for testing for transmission ratio distortion. *Genetics*, *198*(4), 1357–1367. https://doi.org/10.1534/genetics.114.169607

Casellas, J., Id-Lahoucine, S., & Cánovas, A. (2020). Discriminating between allele- and genotype-specific transmission ratio distortion. *Animal Genetics*, *51*(6), 847–854. https://doi.org/10.1111/age.13007

Charlier, C., Agerholm, J. S., Coppieters, W., Karlskov-Mortensen, P., Li, W., de Jong, G., Fasquelle, C., Karim, L., Cirera, S., Cambisano, N., Ahariz, N., Mullaart, E., Georges, M., & Fredholm, M. (2012). A Deletion in the Bovine FANCI Gene Compromises Fertility by Causing Fetal Death and Brachyspina. *PLoS ONE*, *7*(8), e43085. https://doi.org/10.1371/journal.pone.0043085

Chen, S. Y., Schenkel, F. S., Melo, A. L. P., Oliveira, H. R., Pedrosa, V. B., Araujo, A. C., Melka, M. G., & Brito, L. F. (2022). Identifying pleiotropic variants and candidate genes for fertility and reproduction traits in Holstein cattle via association studies based on imputed whole-genome sequence genotypes. *BMC Genomics*, *23*(1), 331. https://doi.org/10.1186/s12864-022-08555-z

Cole, J. B., Wiggans, G. R., Ma, L., Sonstegard, T. S., Lawlor, T. J., Crooker, B. A., Van Tassell, C. P., Yang, J., Wang, S., Matukumalli, L. K., & Da, Y. (2011). Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. *BMC Genomics*, *12*. https://doi.org/10.1186/1471-2164-12-408

Cole, J. B., Null, D. J., & VanRaden, P. M. (2016). Phenotypic and genetic effects of recessive haplotypes on yield, longevity, and fertility. *Journal of Dairy Science*, *99*(9), 7274–7288. https://doi.org/10.3168/jds.2015-10777

Costa, R. B., Camargo, G. M. F., Diaz, I. D. P. S., Irano, N., Dias, M. M., Carvalheiro, R., Boligon, A. A., Baldi, F., Oliveira, H. N., Tonhati, H., & Albuquerque, L. G. (2015). Genome-wide association study of reproductive traits in Nellore heifers using Bayesian inference. *Genetics Selection Evolution*, *47*(1), 67. https://doi.org/10.1186/s12711-015-0146-0

Dias, M. S., Pedrosa, V. B., Rocha da Cruz, V. A., Silva, M. R., & Batista Pinto, L. F. (2024). Genome-wide association and functional annotation analysis for the calving interval in Nellore cattle. *Theriogenology*, *218*, 214–222. https://doi.org/10.1016/j.theriogenology.2024.01.034

Diskin, M. G., & Morris, D. G. (2008). Embryonic and Early Foetal Losses in Cattle and Other Ruminants. *Reproduction in Domestic Animals*, *43*(2), 260–267. https://doi.org/10.1111/j.1439-0531.2008.01171.x

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics* (3rd ed., Vol. 1). Pearson.

Fishman, L., & Mcintosh, M. (2019). Standard Deviations: The Biological Bases of Transmission Ratio Distortion. *Annual Review of Genetics*, *53*(1), 347–372. https://doi.org/10.1146/annurev-genet-112618

Friocourt, G., Perrin, A., Saunders, P. A., Nikalayevich, E., Voisset, C., Coutton, C., Martinez, G., & Morel, F. (2023). Bypassing Mendel's First Law: Transmission Ratio Distortion in Mammals. *International Journal of Molecular Sciences*, *24*(2), 1600. https://doi.org/10.3390/ijms24021600

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. (2013). Detection of Haplotypes Associated with Prenatal Death in Dairy Cattle and Identification of Deleterious Mutations in GART, SHBG and SLC37A2. *PLoS ONE*, *8*(6), e65550. https://doi.org/10.1371/journal.pone.0065550

Gallagher, M. D., & Chen-Plotkin, A. S. (2018). The Post-GWAS Era: From Association to Function. *American Journal of Human Genetics*, *102*(5), 717–730. https://doi.org/10.1016/j.ajhg.2018.04.002

Gershenson, S. (1928). A New Sex-Ratio Abnormality in DROSOPHILA OBSCURA*. *Genetics*, *13*(6), 488. https://academic.oup.com/genetics/article/13/6/488/6063368

Häfliger, I. M., Spengeler, M., Seefried, F. R., & Drögemüller, C. (2022). Four novel candidate causal variants for deficient homozygous haplotypes in Holstein cattle. *Scientific Reports*, *12*(1), 5435. https://doi.org/10.1038/s41598-022-09403-6

Häggman, J., & Uimari, P. (2017). Novel harmful recessive haplotypes for reproductive traits in pigs. *Journal of Animal Breeding and Genetics*, *134*(2), 129–135. https://doi.org/10.1111/jbg.12240

Hartfield, M., & Otto, S. P. (2011). Recombination and hitchhiking of deleterious alleles. *Evolution*, *65*(9), 2421–2434. https://doi.org/10.1111/j.1558-5646.2011.01311.x

Hayes, B. J., Bowman, P. J., Chamberlain, A. J., & Goddard, M. E. (2009). Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science*, *92*(2), 433–443. https://doi.org/10.3168/jds.2008-1646

Hirschhorn, J. N., & Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics*, *6*(2), 95–108. https://doi.org/10.1038/nrg1521

Hoff, J. L., Decker, J. E., Schnabel, R. D., & Taylor, J. F. (2017). Candidate lethal haplotypes and causal mutations in Angus cattle. *BMC Genomics*, *18*(1), 799. https://doi.org/10.1186/s12864-017-4196-2

Hou, L., & Zhao, H. (2013). A review of post-GWAS prioritization approaches. *Frontiers in Genetics*, *4*(1), 280. https://doi.org/10.3389/fgene.2013.00280

Howard, D. M., Pong-Wong, R., Knap, P. W., & Woolliams, J. A. (2017). Use of haplotypes to identify regions harbouring lethal recessive variants in pigs. *Genetics Selection Evolution*, *49*(1), 57. https://doi.org/10.1186/s12711-017-0332-3

Howard, J. T., Pryce, J. E., Baes, C., & Maltecca, C. (2017). Invited review: Inbreeding in the genomics era: Inbreeding, inbreeding depression, and management of

genomic variability. *Journal of Dairy Science*, *100*(8), 6009–6024. https://doi.org/10.3168/jds.2017-12787

Huang, L. O., Labbe, A., & Infante-Rivard, C. (2013). Transmission ratio distortion: Review of concept and implications for genetic association studies. *Human Genetics*, *132*(3), 245–263. https://doi.org/10.1007/s00439-012-1257-0

Id-Lahoucine, S., Cánovas, A., Jaton, C., Miglior, F., Fonseca, P. A. S., Sargolzaei, M., Miller, S., Schenkel, F. S., Medrano, J. F., & Casellas, J. (2019). Implementation of Bayesian methods to identify SNP and haplotype regions with transmission ratio distortion across the whole genome: TRDscan v.1.0. *Journal of Dairy Science*, *102*(4), 3175–3188. https://doi.org/10.3168/jds.2018-15296

Id-Lahoucine, S., Casellas, J., Lu, D., Sargolzaei, M., Miller, S., & Cánovas, A. (2023a). Distortion of Mendelian segregation across the Angus cattle genome uncovering regions affecting reproduction. *Scientific Reports 2023 13:1*, *13*(1), 1–10. https://doi.org/10.1038/s41598-023-37710-z

Id-Lahoucine, S., Casellas, J., Suárez-Vega, A., Fonseca, P. A. S., Schenkel, F. S., Sargolzaei, M., & Cánovas, A. (2023b). Unravelling transmission ratio distortion across the bovine genome: identification of candidate regions for reproduction defects. *BMC Genomics*, *24*(1), 1–16. https://doi.org/10.1186/S12864-023-09455-6/TABLES/4

Irano, N., De Camargo, G. M. F., Costa, R. B., Terakado, A. P. N., Magalhães, A. F. B., Silva, R. M. D. O., Dias, M. M., Bignardi, A. B., Baldi, F., Carvalheiro, R., De Oliveira, H. N., & De Albuquerque, L. G. (2016). Genome-wide association study for indicator traits of sexual precocity in Nellore cattle. *PLoS ONE*, *11*(8), e0159502. https://doi.org/10.1371/journal.pone.0159502

Jenko, J., McClure, M. C., Matthews, D., McClure, J., Johnsson, M., Gorjanc, G., & Hickey, J. M. (2019). 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*, *51*(1), 9. https://doi.org/10.1186/s12711-019-0452-z

Johnson, R., Burch, K. S., Hou, K., Paciuc, M., Pasaniuc, B., & Sankararaman, S. (2021). Estimation of regional polygenicity from GWAS provides insights into the genetic architecture of complex traits. *PLoS Computational Biology*, *17*(10), e1009483. https://doi.org/10.1371/journal.pcbi.1009483

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *28*(1), 27–30. http://www.genome.ad.jp/kegg/

Larracuente, A. M., & Presgraves, D. C. (2012). The selfish Segregation Distorter gene complex of Drosophila melanogaster. *Genetics*, *192*(1), 33–53. https://doi.org/10.1534/genetics.112.141390

Laseca, N., Cánovas, Á., Valera, M., Id-Lahoucine, S., Perdomo-González, D. I., Fonseca, P. A. S., Demyda-Peyrás, S., & Molina, A. (2023). Genomic screening of allelic and genotypic transmission ratio distortion in horse. *PLoS ONE*, *18*(8), e0289066. https://doi.org/10.1371/journal.pone.0289066

Liu, H., Sørensen, A. C., Meuwissen, T. H. E., & Berg, P. (2014). Allele frequency changes due to hitch-hiking in genomic selection programs. *Genetics Selection Evolution*, *46*(1), 8. https://doi.org/10.1186/1297-9686-46-8

Lozada-Soto, E. A., Maltecca, C., Lu, D., Miller, S., Cole, J. B., & Tiezzi, F. (2021). Trends in genetic diversity and the effect of inbreeding in American Angus cattle under genomic selection. *Genetics Selection Evolution*, *53*(1), 50. https://doi.org/10.1186/s12711-021-00644-z

Lucy, M. C. (2019). Symposium review: Selection for fertility in the modern dairy cow—Current status and future direction for genetic selection. *Journal of Dairy Science*, *102*(4), 3706–3721. https://doi.org/10.3168/jds.2018-15544

Makanjuola, B. O., Miglior, F., Abdalla, E. A., Maltecca, C., Schenkel, F. S., & Baes, C. F. (2020). Effect of genomic selection on rate of inbreeding and coancestry and effective population size of Holstein and Jersey cattle populations. *Journal of Dairy Science*, *103*(6), 5183–5199. https://doi.org/10.3168/jds.2019-18013

Melo, T. P. de, De Camargo, G. M. F., De Albuquerque, L. G., & Carvalheiro, R. (2017). Genome-wide association study provides strong evidence of genes affecting the reproductive performance of Nellore beef cows. *PLoS ONE*, *12*(5). https://doi.org/10.1371/journal.pone.0178551

Merrick, L. F., Burke, A. B., Zhang, Z., & Carter, A. H. (2022). Comparison of Single-Trait and Multi-Trait Genome-Wide Association Models and Inclusion of Correlated Traits in the Dissection of the Genetic Architecture of a Complex Trait in a Breeding Program. *Frontiers in Plant Science*, *12*, 772907. https://doi.org/10.3389/fpls.2021.772907

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics*, *157*(4), 1819. https://academic.oup.com/genetics/article/157/4/1819/6048353

Misztal, I., & Lourenco, D. (2024). Potential negative effects of genomic selection. *Journal of Animal Science*, *102*, skae155. https://doi.org/10.1093/jas/skae155

Mulim, H. A., Campos, G. S., Cardoso, F. F., & Rojas de Oliveira, H. (2025). Exploring inbreeding depression in Brazilian Angus cattle population using pedigree and genomic data. *Frontiers in Genetics*, *16*(1), 1–10. https://doi.org/10.3389/fgene.2025.1613820

Ogunbawo, A. R., Hidalgo, J., Mulim, H. A., Carrara, E. R., Ventura, H. T., Souza, N. O., Lourenco, D., & Oliveira, H. R. (2025). Applying the algorithm for Proven and young in GWAS Reveals high polygenicity for key traits in Nellore cattle. *Frontiers in Genetics*, *16*(1), 1–29. https://doi.org/10.3389/fgene.2025.1549284

Olasege, B. S., Tahir, M. S., Gouveia, G. C., Kour, J., Porto-Neto, L. R., Hayes, B. J., & Fortes, M. R. S. (2021). Genetic parameter estimates for male and female fertility traits using genomic data to improve fertility in Australian beef cattle. *Animal Production Science*, *61*(18), 1863–1872. https://doi.org/10.1071/AN21097

Parker Gaddis, K. L., Null, D. J., & Cole, J. B. (2016). Explorations in genome-wide association studies and network analyses with dairy cattle fertility traits. *Journal of Dairy Science*, *99*(8), 6420–6435. https://doi.org/10.3168/jds.2015-10444

Pryce, J. E., Bolormaa, S., Chamberlain, A. J., Bowman, P. J., Savin, K., Goddard, M. E., & Hayes, B. J. (2010). A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using variable length haplotypes. *Journal of Dairy Science*, *93*(7), 3331–3345. https://doi.org/10.3168/jds.2009-2893

Rodrigues, G. R. D., Brito, L. F., Mota, L. F. M., Cyrillo, J. N. S. G., Valente, J. P. S., Benfica, L. F., Silva Neto, J. B., Borges, M. S., Monteiro, F. M., Faro, L. El, Albuquerque, L. G., & Mercadante, M. E. Z. (2024). Genome-wide association studies and functional annotation of pre-weaning calf mortality

and reproductive traits in Nellore cattle from experimental selection lines. *BMC Genomics*, *25*(1), 1196. https://doi.org/10.1186/s12864-024-11113-4

Rodrigues, G. R. D., Brito, L. F., Mota, L. F. M., Cyrillo, J. N. S. G., Silva Neto, J. B., Campos, M. A. F., El Faro, L., Albuquerque, L. G., & Mercadante, M. E. Z. (2025b). Unraveling genomic regions with transmission ratio distortion harboring putative lethal alleles and their biological implications in Nellore cattle from experimental selection lines. *Journal of Animal Science*, *103*(1), skaf208. https://doi.org/10.1093/jas/skaf208

Rodrigues, G. R. D., Cyrillo, J. N. S. G., Mota, L. F. M., Schmidt, P. I., Valente, J. P. S., Oliveira, E. S., Albuquerque, L. G., Brito, L. F., & Mercadante, M. E. Z. (2025a). Effect of genomic regions harboring putative lethal haplotypes on reproductive performance in closed experimental selection lines of Nellore cattle. *Scientific Reports*, *15*(1), 4113. https://doi.org/10.1038/s41598-025-88501-7

Sandler, L., & Novitski, E. (1957). Meiotic Drive as an Evolutionary Force. *The American Naturalist*, *91*(857), 105–110. https://www.jstor.org/stable/2458433

Sbardella, A. P., Watanabe, R. N., Costa, R. M. da, Bernardes, P. A., Braga, L. G., Rey, F. S. B., Lôbo, R. B., & Munari, D. P. (2021). Genome-wide association study provides insights into important genes for reproductive traits in nelore cattle. *Animals*, *11*(5), 1386. https://doi.org/10.3390/ani11051386

Schaeffer, L. R. (2006). Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics*, *123*(4), 218–223.

Schmidt, P. I., Mota, L. F. M., Fonseca, L. F. S., dos Santos Silva, D. B., Frezarim, G. B., Arikawa, L. M., de Abreu Santos, D. J., Magalhães, A. F. B., Cole, J. B., Carvalheiro, R., de Oliveira, H. N., Null, D. J., VanRaden, P., Ma, L., & de Albuquerque, L. G. (2023). Identification of candidate lethal haplotypes and genomic association with post-natal mortality and reproductive traits in Nellore cattle. *Scientific Reports*, *13*(1), 1–12. https://doi.org/10.1038/s41598-023-37586-z

Silberstein, M., Nesbit, N., Cai, J., & Lee, P. H. (2021). Pathway analysis for genome-wide genetic variation data: Analytic principles, latest developments, and new opportunities. In *Journal of Genetics and Genomics* (Vol. 48, Issue 3, pp. 173–183). Institute of Genetics and Developmental Biology. https://doi.org/10.1016/j.jgg.2021.01.007

Speidel, S. E., Buckley, B. A., Boldt, R. J., Enns, R. M., Lee, J., Spangler, M. L., & Thomas, M. G. (2018). Genome-wide association study of stayability and heifer pregnancy in red angus cattle. *Journal of Animal Science*, *96*(3), 846–853. https://doi.org/10.1093/jas/sky041

Stegemiller, M. R., Murdoch, G. K., Rowan, T. N., Davenport, K. M., Becker, G. M., Hall, J. B., & Murdoch, B. M. (2021). Genome-wide association analyses of fertility traits in beef heifers. *Genes*, *12*(2), 1–15. https://doi.org/10.3390/genes12020217

Strome, S., Bhalla, N., Kamakaka, R., Sharma, U., & Sullivan, W. (2024). Clarifying Mendelian vs non-Mendelian inheritance. *Genetics*, *227*(3), iyae078. https://doi.org/10.1093/genetics/iyae078

Sweett, H., Fonseca, P. A. S., Suárez-Vega, A., Livernois, A., Miglior, F., & Cánovas, A. (2020). Genome-wide association study to identify genomic regions and positional candidate genes associated with male fertility in beef cattle. *Scientific Reports*, *10*(1), 20102. https://doi.org/10.1038/s41598-020-75758-3

Thomsen, B., Horn, P., Panitz, F., Bendixen, E., Petersen, A. H., Holm, L. E., Nielsen, V. H., Agerholm, J. S., Arnbjerg, J., & Bendixen, C. (2006). A missense mutation in the bovine SLC35A3 gene, encoding a UDP-N-acetylglucosamine transporter, causes complex vertebral malformation. *Genome Research*, *16*(1), 97–105. https://doi.org/10.1101/gr.3690506

Toghiani, S., Hay, E., Sumreddee, P., Geary, T. W., Rekaya, R., & Roberts, A. J. (2017). Genomic prediction of continuous and binary fertility traits of females in a composite beef cattle breed. *Journal of Animal Science*, *95*(11), 4787–4795. https://doi.org/10.2527/jas2017.1944

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., & Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews Methods Primers*, *1*(1), 59. https://doi.org/10.1038/s43586-021-00056-9

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

VanRaden, P. M., Olson, K. M., Null, D. J., & Hutchison, J. L. (2011). Harmful recessive effects on fertility detected by absence of homozygous haplotypes.

Journal of Dairy Science, *94*(12), 6153–6161. https://doi.org/10.3168/jds.2011-4624

Wang, H., Misztal, I., Aguilar, I., Legarra, A., & Muir, W. M. (2012). Genome-wide association mapping including phenotypes from relatives without genotypes. *Genetics Research*, *94*(2), 73–83. https://doi.org/10.1017/S0016672312000274

Wu, X., Mesbah-Uddin, M., Guldbrandtsen, B., Lund, M. S., & Sahana, G. (2019). Haplotypes responsible for early embryonic lethality detected in Nordic Holsteins. *Journal of Dairy Science*, *102*(12), 11116–11123. https://doi.org/10.3168/jds.2019-16651

CHAPTER 2: GENOME-WIDE ASSOCIATION STUDIES AND FUNCTIONAL ANNOTATION OF PRE-WEANING CALF MORTALITY AND REPRODUCTIVE TRAITS IN NELLORE CATTLE FROM EXPERIMENTAL SELECTION LINES

Gustavo R. D. Rodrigues[1,2,3,*], Luiz F. Brito[2], Lúcio F. M. Mota[1], Joslaine N. S. G. Cyrillo[3], Júlia P. S. Valente[1,3], Lorena F. Benfica[2], João B. Silva Neto[1,2], Marcelo S. Borges[1,3], Fábio M. Monteiro[1,3], Lenira El Faro[1,3], Lucia G. Albuquerque[1,4], Maria E. Z. Mercadante[1,3,4,*]

[1]São Paulo State University, School of Agriculture and Veterinary Science, Jaboticabal, SP, 14884-900, Brazil
[2]Department of Animal Sciences, Purdue University, West Lafayette, IN, 47907, USA
[3]Institute of Animal Science, Beef Cattle Research Center, Sertãozinho, SP, 14174-000, Brazil
[4]National Council for Science and Technological Development, Brasilia, DF, 71605-001, Brazil

Corresponding authors: Gustavo R. D. Rodrigues: grd.rodrigues@unesp.br and Maria E. Z. Mercadante: mezmercadante@gmail.com

Published in BMC Genomics (ISSN: 1471-2164)

Received: 24 October 2024
Accepted: 02 December 2024
Published: 18 December 2024

Cite this article: Rodrigues, G.R.D., Brito, L.F., Mota, L.F.M. *et al.* Genome-wide association studies and functional annotation of pre-weaning calf mortality and reproductive traits in Nellore cattle from experimental selection lines. *BMC Genomics* **25**, 1196 (2024). https://doi.org/10.1186/s12864-024-11113-4

Abstract

Background

Reproductive efficiency is crucial for the long-term economic sustainability of beef cattle production. Pregnancy loss and stillbirth are complex reproductive traits that do not yet have their genomic background fully understood, especially in zebu breeds (*Bos taurus indicus*). Hence, this study aimed to perform a genome-wide association study (GWAS) and functional annotation for conception success (CS), pregnancy loss (PL), stillbirth (SB), and pre-weaning calf mortality (PWM) in Nellore cattle. In this study, 3,728 cows with 17,094 reproductive records and 11,785 calves were evaluated. A total of 3,351 genotyped animals and 383,739 SNP markers were considered for GWAS analyses. SNP effects were estimated using the weighted single-step GWAS (WssGWAS), which considered two iterations. The top ten genomic windows with the highest contribution to the additive genetic variance of the traits were selected for gene annotation. Candidate genes were then analyzed for Gene Ontology terms (GO) and metabolic pathways.

Results

The top ten genomic windows that explained the largest proportion of the direct additive genetic variance (σ^2_a) for CS, PL, SB, and PWM accounted for 17.03% (overlapping with 79 genes), 16.76% (57 genes), 11.71% (73 genes), and 12.03% (65 genes) of the total σ^2_a, respectively. For CS, significant GO terms included Somitogenesis (GO:0001756), Somite Development (GO:0061053), and Chromosome Segregation (GO:0007059). Considering PL, the processes annotated were the Regulation of Hormone Secretion (GO:0046883), and Hormone Transport (GO:0009914), along with the Glucagon Signaling Pathway (bta04922). Embryonic Development (GO:0045995), and Cerebellum Development (GO:0021549) were the main biological processes found in the gene enrichment analysis for SB. For PWM, the Regulation of Glucose metabolic processes (GO:0010906), Zinc Ion Homeostasis (GO:0055069), Lactation (GO:0007595), and Regulation of Insulin Secretion (GO:0050796) were the most significant GO terms observed.

Conclusions

These findings provide valuable information on genomic regions, candidate genes, biological processes, and metabolic pathways that may significantly influence the expression of complex reproductive traits in Nellore cattle, offering potential contributions to breeding strategies and future genomic selection strategies.

Keywords: beef cattle, pregnancy loss, reproduction, single nucleotide polymorphism, stillbirth.

2.1. Introduction

Reproductive efficiency holds considerable relevance for the long-term economic sustainability of beef cattle production. Reproductive traits are directly related to the annual number of calves born, and thus, a low pregnancy rate or a high pre-weaning mortality rate can compromise the number of animals available for sale and breeding. This situation can challenge the long-term viability of beef cattle production, as most of the beef industry income is derived from the sale of animals for meat production [1–3].

Some reproductive traits are difficult to measure in commercial herds and/or are typically recorded late in the life of dams (e.g., pregnancy loss, embryo loss, and stillbirth), limiting their potential for inclusion in animal breeding programs [4]. Likewise, the environment exerts a significant influence on the phenotypic expression of these traits, which reduces heritability estimates and genetic gain due to selection [5]. Thus, understanding the genetic background of reproductive traits in beef cattle remains a research area of great interest when optimizing breeding schemes [6]. Genome-wide association studies (GWAS) are commonly used to identify genomic regions and genes potentially associated with relevant traits across species [6–8]. By identifying these genomic regions, it is possible to annotate genes and biological or metabolic pathways that regulate the phenotypic expression of complex traits.

Pregnancy loss and stillbirth represent complex phenomena with incomplete knowledge about their genetic determinants, particularly in Zebu cattle breeds. Notably, there is a significant gap in research specifically targeting the identification of genes, biological processes, and metabolic pathways linked to the manifestation of these traits in Nellore cattle (*Bos taurus indicus*). To date, no studies have comprehensively explored the genomic architecture underlying these reproductive traits in this breed, underscoring the need for genetic evaluations. Furthermore, the population sampled for this study is resultant from a Nellore experimental breeding program initiated in 1980, in which various complex traits are collected on individual animals from birth to adult age [9–12]. This systematic collection of information enables the genomic analysis of traits that are not typically evaluated in conventional

animal breeding programs, such as reproductive traits measured longitudinally and expressed in the adult life of cows. Hence, the objectives of this study were to 1) perform a GWAS for conception success (CS), pregnancy loss (PL), stillbirth (SB), and pre-weaning mortality (PWM) in Nellore cattle, and 2) annotate candidate genes, biological processes, and metabolic pathways associated with the phenotypic expression of these traits.

2.2. Materials and Methods

2.2.1. Ethical statement

Data was obtained from an existing database, and therefore, approval from the Ethics Committee was not required.

2.2.2. Phenotypic data

All datasets used in this study are from a Nellore cattle population from an experimental selection design belonging to the Beef Cattle Research Center (Institute of Animal Science – IZ, Sertãozinho, SP, Brazil). The herd is divided into three selection lines that have been selected for 44 years, considering different criteria for yearling weight over the decades. More details about the Nellore IZ experimental breeding program and the selection lines can be found in Benfica et al. [13] and Rodrigues et al. [11]. In this study, data from 3,728 cows (17,094 reproductive records) and 11,785 calves were used for the analyses. Data was collected from breeding seasons that occurred between 1980 and 2022. The traits analyzed were CS, PL, SB, and PWM.

Each year, cows and heifers were subjected to either natural mating or fixed-time artificial insemination during a 90-day breeding season. Approximately 30 days after the end of the breeding season, ultrasound examinations were conducted to determine CS. Females were categorized as either pregnant (1 = conception success) or non-pregnant (0 = conception failure). If a female was diagnosed as pregnant but did not give birth to a calf, it was classified as a PL (assigned as 1); otherwise, females that successfully gave birth were classified as non-pregnancy loss (0). The PL trait began to be measured only from 2005 onward, which explains the smaller number of observations compared to the other traits evaluated in this study (Table 1).

For cows whose calves died within the first 48 hours after birth, the event was classified as an SB (assigned as 1). Conversely, if the calves survived beyond this period, the event was classified as non-stillborn (assigned as 0). Calves were weaned at around 7 months. PWM was determined by assigning a value of 1 (1 = mortality) for calves that did not reach weaning age and 0 (0 = survived) for animals that survived until weaning. As this is an experimental breeding program, the reproductive records used for this study were systematically controlled, providing detailed information about the birth and death dates of each calf.

Table 1 Descriptive statistics for the traits evaluated in this study.

Trait	N_{obs}[5]	N_{anim}[6]	CG[7]	Freq (%)[8]	h^2 (± SD)[9]	h^2_m (± SD)[10]	t (± SD)[11]
CS[1]	17,094	3,728	126	74.68	0.07 ± 0.01	-	0.07 ± 0.01
PL[2]	5,193	1,608	51	5.93	0.06 ± 0.03	-	0.10 ± 0.04
SB[3]	12,102	2,941	126	2.62	0.15 ± 0.03	-	0.18 ± 0.04
PWM[4]	11,785	11,785	252	10.33	0.11 ± 0.03	0.08 ± 0.02	-

[1]CS: conception success; [2]PL: pregnancy loss; [3]SB: stillbirth; [4]PWM: pre-weaning mortality; [5]N_{obs}: number of observations; [6]N_{anim}: number of animals; [7]CG: number of contemporary groups; [8]Freq: proportion of total observations in which the event was observed (in percentage); [9]h^2: direct heritability; [10]h^2_m: maternal heritability; [11]t: repeatability

2.2.3. Genomic and pedigree data

The pedigree file used for the analyses included 13,088 individuals, 486 sires, and 3,007 dams. The genomic dataset included 3,351 animals, in which 1,766 males and 1,581 females were genotyped with different SNP chip assays (Supplementary File 1, Table S1). Animals (n = 2,571) genotyped with medium-density SNP panels (50K and 75K SNP panels) were imputed to the Illumina BovineHD panel (770K) using the FImpute v.3 software [14] considering a reference population of 6,862 animals genotyped with the HD SNP chip. The genome coordinates considered in the genotype imputation analyses are from ARS-UCD1.2 *Bos taurus* genome assembly [15]. The expected genotype imputation accuracy was higher than 0.97 [16]. After genotype imputation, 612,154 autosomal SNP markers remained in the genomic dataset. The software PreGSf90 [17] was used to perform quality control (QC) of SNP markers and

samples. The criteria considered for genotype QC were: 1) removal of SNPs with minor allele frequency (MAF) < 0.05; 2) removal of SNPs and samples with a call rate < 0.90; 3) removal of SNPs with extreme deviations (p-value ≤ 10^{-5}) from Hardy-Weinberg equilibrium; 4) removal of SNPs with unknown or duplicated genomic positions; and 5) removal of SNPs or samples with Mendelian conflicts. After genomic quality control, 3,351 animals and 383,739 SNP markers remained for further analyses.

2.2.4. Genome-wide association analyses

The contemporary groups (CG) for CS, PL, and SB were formed by the combination of the breeding year (from 1980 to 2022) and selection line (1, 2, and 3). Season information was not included in the definition of CGs because all cows were matted during a single breeding season. For PWM, the CG was defined considering the calves' birth year (1981 – 2023), selection line (1, 2, and 3), and sex (male or female). CGs with less than 5 animals and showing no phenotypic variability within CG were excluded from further analyses. Table 1 presents the descriptive statistics for all the traits evaluated in this study.

For the genetic analyses of the traits CS, PL, and SB, the effects considered were the CGs (systematic effect), cows' age at breeding (covariate with linear and quadratic effects), direct additive genetic and permanent environmental effects (random effects). The model in matrix denotation can be described as:

$$\mathbf{l} = \mathbf{X\beta} + \mathbf{Z_1 a} + \mathbf{Z_2 pe} + \mathbf{e}$$

where **l** represents the linear predictor for binary traits; $\boldsymbol{\beta}$ is the vector with solutions for the systematic effects and covariates; $\mathbf{a}$ is the vector with the solutions for the direct additive genetic effect, assumed as $\mathbf{a} \sim N(0, \mathbf{H}\sigma^2_{\mathbf{a}})$, where **H** is the relationship matrix that combines pedigree information (**A**) and genomic (**G**) relationships, and $\sigma^2_{\mathbf{a}}$ is the direct additive genetic variance; **pe** is the vector with the solutions for the permanent environmental effect, assumed as $\mathbf{pe} \sim N(0, \mathbf{I}\sigma^2_{\mathbf{pe}})$, where **I** is an identity matrix and $\sigma^2_{\mathbf{pe}}$ is the permanent environment variance; $\mathbf{e}$ is the vector of residuals effects, assumed as $\mathbf{e} \sim N(0, \mathbf{I}\sigma^2_{\mathbf{e}})$, where $\sigma^2_{\mathbf{e}}$ is the residual variance; **X** is the incidence matrix that relates $\boldsymbol{\beta}$ and **l**; $\mathbf{Z_1}$ is the incidence matrix that relates **a** and **l**; and $\mathbf{Z_2}$ is the incidence matrix that relates **pe** and **l**.

For PWM, the systematic effects considered were the CGs and month of birth of calves (September, October, or November), the covariate included was the cows' age at calving (with linear and quadratic effects), and the random effects included the direct additive genetic, maternal additive genetic, and maternal permanent environmental effects. In matrix denotation this model can be described as:

$$\mathbf{l} = \mathbf{X\beta} + \mathbf{Z_1 a} + \mathbf{Z_3 m} + \mathbf{Z_4 mpe} + \mathbf{e}$$

where **l**, $\mathbf{X\beta}$, $\mathbf{Z_1 a}$, and **e** was previously described; **m** is the vector with the solutions for the maternal additive genetic effect, assumed as **m** ~ N(0, $\mathbf{H\sigma^2_m}$), where $\boldsymbol{\sigma}^2_{\mathbf{m}}$ is the maternal additive genetic variance; **mpe** is the vector with the solutions for the maternal permanent environmental effect, assumed as **mpe** ~ N(0, $\mathbf{I\sigma^2_{mpe}}$), where $\boldsymbol{\sigma}^2_{\mathbf{mpe}}$ is the maternal permanent environmental variance; $\mathbf{Z_3}$ is the incidence matrix that relates **m** and **l**; and $\mathbf{Z_4}$ is the incidence matrix that relates **mpe** and **l**. The direct additive genetic and maternal additive genetic effects were considered to be correlated based on Cov(**a**, **m**) = $\mathbf{H\sigma_{am}}$, where $\boldsymbol{\sigma}_{\mathbf{am}}$ is the covariance between **a** and **m**.

All traits were evaluated considering the single-step Genomic Best Linear Unbiased Prediction (ssGBLUP) method via a univariate threshold animal model to estimate the variance components via Bayesian inference. An underlying distribution was assumed for all traits as follows:

$$\mathbf{f(y|l_i)} = \prod_{\mathbf{i=1}}^{\mathbf{n_i}} \mathbf{1(l_i < t_i)1(y = 0) + 1(l_i > t_i)1(y = 1)}$$

where **y** is the binary trait (CS, PL, SB, or PWM), $\mathbf{l_i}$ is the underlying liability of observation *i*; $\mathbf{t_i}$ is the threshold that defines the category response for the traits, and $\mathbf{n_i}$ corresponds to the number of observations.

In the ssGBLUP method, the inverse of the pedigree-based relationship matrix $\mathbf{A^{-1}}$ is replaced by the inverse of a hybrid pedigree-genomic relationship matrix $\mathbf{H^{-1}}$, as described by Aguilar et al. [18]

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

where $\mathbf{A_{22}^{-1}}$ is the inverse of the relationship matrix for the genotyped animals and $\mathbf{G^{-1}}$ is the inverse of the genomic relationship matrix described by VanRaden [19]:

$$\mathbf{G} = \frac{\mathbf{ZZ'}}{\mathbf{2\sum_{i=1}^{m} p_i(1 - p_i)}}$$

where **Z = (M – P)**, in which **M** is the SNP incidence matrix, with *m* columns (number of SNP markers) and *n* lines (number of genotyped animals). The elements in **M** were set to 0, 1, and 2 for the genotypes AA, AB, and BB, respectively. **P** is the matrix with the allele frequencies expressed as 2pi; and p_i is the frequency of the *i*th SNP marker.

Genetic analyses were performed in the gibbsf90+ and postgibbsf90+ software [17] to obtain the posterior distributions of genetic parameters. The Gibbs sampling consisted of a chain of 500,000 cycles, with a burn-in of the first 150,000 iterations and estimates stored every ten cycles, totaling 35,000 samples for the inferences about the variance components and genetic parameters (heritability and repeatability). The convergence of the estimates obtained by the Monte Carlo Markov Chain method was checked using the Geweke test [20] and visual inspection through the BOA package available in the R software [21].

The SNP effects were calculated in an iterative method based on the weighted single-step GWAS (WssGWAS) procedure [22]. This analysis used the genomic estimated breeding values (GEBV) of the genotyped animals [22] and was performed in the postGSf90 software [17]. The equation for obtaining the effect of the SNPs can be described as:

$$\hat{\mathbf{u}} = \lambda \mathbf{D}\mathbf{Z}'\mathbf{G}^{*-1}\hat{\mathbf{a}}_{\mathbf{g}} = \mathbf{D}\mathbf{Z}'[\mathbf{Z}\mathbf{D}\mathbf{Z}']^{-1}\hat{\mathbf{a}}_{\mathbf{g}}$$

where $\hat{\mathbf{u}}$ is a vector with the effects of each SNP; **D** is a diagonal matrix with the weights for the effect of each SNP; **Z** is the matrix that relates the genotypes of each locus; $\mathbf{G}^{*-1}$ is the inverse matrix of weighted genomic relationships; $\hat{\mathbf{a}}_{\mathbf{g}}$ is the vector with the predicted GEBV, which is represented by a function of the effects of SNP ($\hat{\mathbf{a}}_{\mathbf{g}}$ **= Zu**); λ is a weighting factor based on SNP frequencies ($\frac{1}{2\sum_{i=1}^{m} p_i(1-p_i)}$).

The effects of each SNP on the total additive genetic variance were estimated considering two iterations. In the first iteration, **D** is an identity matrix (**D** = **I**), and in the second iteration, **D** is a diagonal matrix with the weights for the SNP markers computed in the first step. The proportion of the genetic variance explained by windows of 1 Mb was calculated according to Wang et al. [22]:

$$\frac{\mathrm{Var}(\mathbf{a_i})}{\boldsymbol{\sigma}^2_{\mathbf{a}}} \times 100\% = \frac{\mathrm{Var}\left(\sum_{j=1} \mathbf{z_j}\hat{\mathbf{u}}_{\mathbf{j}}\right)}{\boldsymbol{\sigma}^2_{\mathbf{a}}} \times 100\%$$

where $\mathbf{a_i}$ is the genetic value of the i[th] region of 1 Mb; σ^2_a is the genetic variance (direct additive genetic and maternal additive genetic); $\mathbf{z_j}$ is the vector with the genotype of

the j^{th} SNP for all animals; and $\hat{u}_j$ is the estimated effect for the j^{th} SNP within the i-th region.

2.2.5. In-silico functional genomic analyses

The genetic variance explained by the genomic windows were used to construct Manhattan plots using the CMplot v4.3.0 package in R [23]. The top ten genomic window regions with the higher contribution for the direct additive genetic variance (CS, PL, SB, and PWM) and maternal additive genetic variance (PWM) were used to annotate genes considering the *Bos taurus* ARS-UCD1.2 assembly [15] as the reference genome. Candidate genes were identified using the BioMart tool from the ENSEMBL software (www.ensembl.org/biomart/martview/).

There is still no consensus on the threshold value for identifying genomic windows associated with traits of interest in GWAS analyses. Silva Neto et al. [24], for instance, established a threshold for genomic windows explaining more than 0.5% of the additive genetic variance, while Carvalho et al. [25] employed a threshold of over 1.0%. Both studies focused on growth traits of Nellore cattle. Alternatively, other researchers have adopted a different approach by selecting the top 10 genomic windows that account for the largest proportion of total genetic variance for each trait. This method has been employed by Irano et al. [26] and Mota et al. [27] for traits related to sexual precocity in Nellore cattle, Magalhães et al. [28] for meat quality traits in Nellore cattle, and Arikawa et al. [29] for carcass traits in Nellore cattle. In light of the ongoing lack of consensus regarding a standard threshold for genomic window selection, we chose to report the top 10 genomic windows in this study. These genomic windows were defined as 1 Mb intervals that accounted for the largest proportion of total additive genetic variance of CS, PL, SB, and PWM, respectively. This approach is consistent with previous studies in Nellore cattle [26-29], which focused on identifying the most informative genomic regions associated with the genetic variance of relevant traits, annotation of candidate genes, and in-silico functional genomic analyses.

Functional classification of genes for biological processes (Gene Ontology - GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were identified using the cluster-Profiler R Package [30].

2.3. Results

2.3.1. Genome-wide association for conception success

The top 10 genomic windows explained 17.03% direct additive genetic variance (σ_a^2) for CS (Table 2 and Figure 1) and were located on 9 *Bos taurus* autosomes (BTA), (1, 4, 5, 7, 14, 19, 25, 26, and 27). Considering the locations of the top 10 genomic windows a total of 79 genes were mapped, including 3 MicroRNA, 71 protein-coding, and 5 snRNA genes. The interactions between the protein-coding genes found for CS are presented in Supplementary File 2, Figure S1.

The genomic window located in BTA 7 (3.03 – 4.03 Mb) explained 3.37% of the σ_a^2 of CS and overlapped with twenty genes (Table 2). The second genomic region that explained most of the σ_a^2 for CS was in BTA27 (10.28 – 11.27 Mb), accounting for 2.61% of the σ_a^2 of this trait, and only the Small Nucleolar RNA, C/D Box 22 (*SNORD22)* gene was annotated in this region. The functional analyses revealed significant GO terms associated with CS considering the gene network (Table 3). The protein kinase, DNA-activated, catalytic subunit (*PRKDC*, BTA14) and WNT family member A3 (*WNT3A*, BTA7) genes were involved in the processes Somitogenesis (GO:0001756) and Somite Development (GO:0061053). The Chromosome Segregation (GO:0007059) with the genes MAU2 sister chromatid cohesion factor (*MAU2*, BTA7), baculoviral IAP repeat containing 5 (*BIRC5*, BTA19), MIS18 kinetochore protein A (*MIS18A*, BTA1), and synaptonemal complex central element protein 1 (*SYCE1*, BTA26) was also associated with CS. No metabolic pathways in the KEGG enrichment analyses were statistically significant for CS (P-value > 0.05).

Table 2 List of the top 10 genomic windows that explained the largest proportion of the direct additive genetic variance (σ_a^2) for conception success (CS) in Nellore cattle.

BTA[1]	Location (Mb)	Genes[2]	σ_a^2 (%)
7	3.03 – 4.03	*WNT3A, WNT9A, PRSS38, SNAP47, JMJD4, ZNF354B, ATP13A1, GMIP, LPAR2, PBX4, U6, CILP2, YJEFN3, TSSK6, GATAD2A, MAU2, SUGP1, TM6SE2, HAPLN4, NCAN*	3.37
27	10.28 – 11.27	*SNORD22*	2.61
25	5.06 – 6.06	*RBFOX1, U6*	2.41
5	117.21 – 118.21	*TBC1D22A, bta-mir-2285o-5*	1.92
19	53.47 – 54.46	*TIMP2, USP36, CYTH1, DNAH17, PGS1, SOCS3, THA1, TMEM235, BIRC5, AFMID, TK1, SYNGR2, TMC8, TMC6, TNRC6C*	1.66
4	11.14 – 12.14	*TFPI2, GNGT1, GNG11, BET1, COL1A2, CASD1, SGCE, PEG10, U6*	1.39
14	18.76 – 19.75	*CEBPD, SPIDR, PRKDC, MCM4, UBE2V2*	1.10
26	49.39 – 50.38	*GLRX3, bta-mir-2397, TCERG1L, SYCE1*	0.89
7	94.11 – 95.10	*Bta-mir-2284z-7, FAM81B, SKIC3, ARSK, RFESD, SPATA9, RHOBTB3*	0.86
1	2.29 – 3.29	*IL10RB, IFNAR2, OLIG1, OLIG2, EPCIP, PAXBP1, SYNJ1, CFAP298, EVA1C, URB1, MRAP, MIS18A*	0.82

[1]BTA: *Bos taurus* autosome; [2]All genes listed in this table were included in the enrichment analyses for this trait.

Table 3 Significant Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses.

Trait	Description	P-value	Gene ID
CS[1]	GO:0001756 - Somitogenesis	0.0217	*PRKDC, WNT3A*
CS[1]	GO:0061053 - Somite development	0.0354	*PRKDC, WNT3A*
CS[1]	GO:0007059 - Chromosome segregation	0.0424	*MAU2, BIRC5, MIS18A, SYCE1*
PL[2]	GO:0046883 - Regulation of hormone secretion	0.0161	*C1QTNF3, RAB11FIP5, PFKM*
PL[2]	GO:0009914 - Hormone transport	0.0306	*C1QTNF3, RAB11FIP5, PFKM*
PL[2]	bta04922 - Glucagon signaling pathway	0.0217	*PFKM, CALM2*
SB[3]	GO:0045995 - Regulation of embryonic development	0.0283	*NOCT, DLL1*
SB[3]	GO:0021549 - Cerebellum development	0.0349	*DLL1, SSTR2*
PWM[4]	GO:0010906 - Regulation of glucose metabolic process	0.0052	*FOXO1, PRKN, USP7*
PWM[4]	GO:0055069 - Zinc ion homeostasis	0.0084	*PRKN, ATP7B*
PWM[4]	GO:0007595 - Lactation	0.0088	*NME1, ATP7B*
PWM[4]	GO:0050796 - Regulation of insulin secretion	0.0129	*ABAT, FOXO1, PRKN*

[1]CS: conception success; [2]PL: pregnancy loss; [3]SB: stillbirth; [4]PWM: pre-weaning mortality.

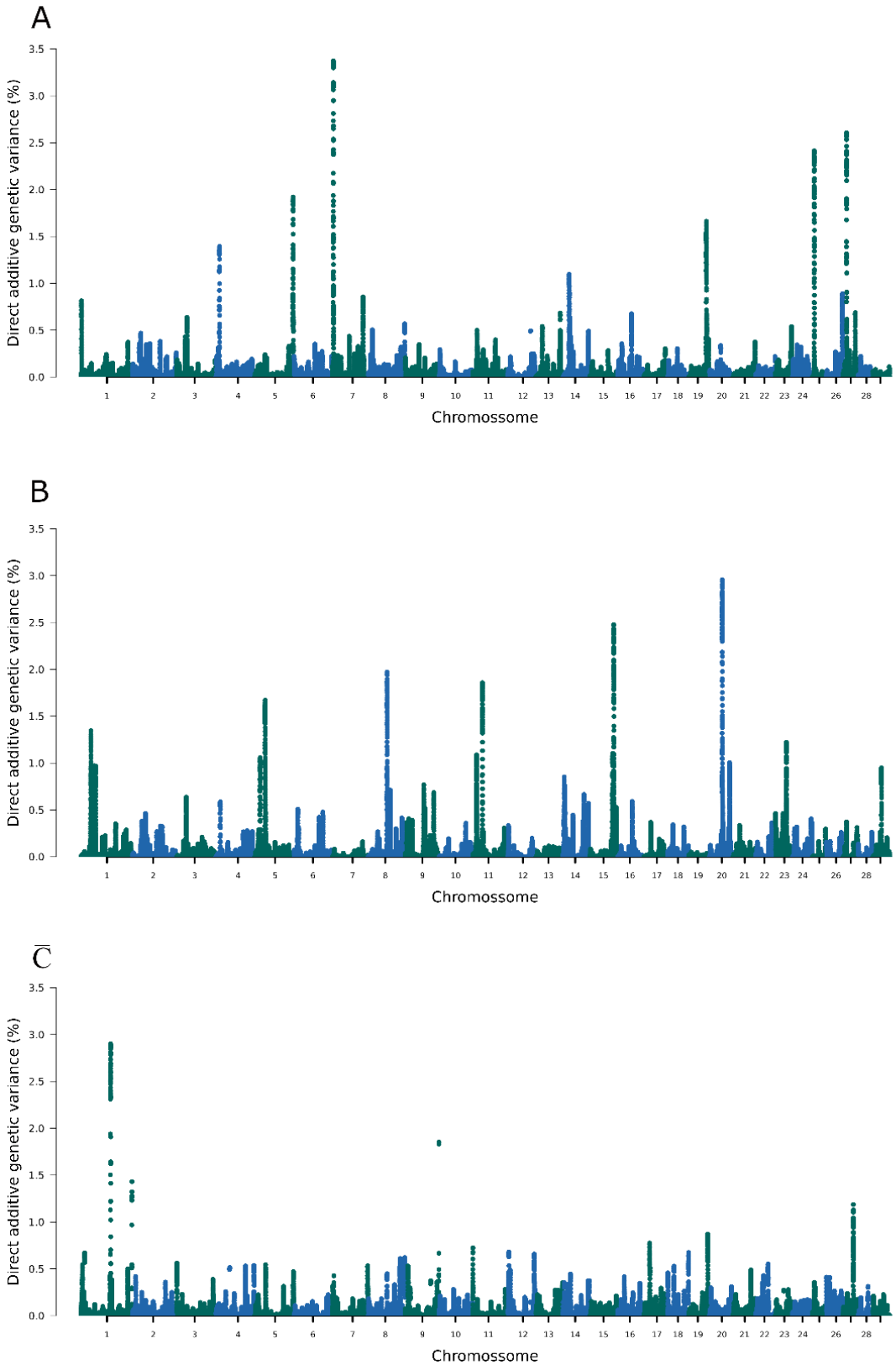

Figure 1 Manhattan plots for the proportion of the total additive genetic variance explained by 1 Mb genomic windows for conception success (A), pregnancy loss (B), and stillbirth (C) in Nellore cattle.

2.3.2. Genome-wide association for pregnancy loss

Regarding PL, the top ten genomic windows associated with this trait accounted for 16.76% of the total σ_a^2 (Table 4 and Figure 1). These genomic windows were located in BTA1, BTA5, BTA8, BTA11, BTA15, BTA20, and BTA23, with 57 genes annotated in these locations, being 3 MicroRNA, 49 protein-coding, and 5 snRNA genes. The interactions between the protein-coding genes are presented in Supplementary file 3, Figure S2.

The genomic window located in BTA20 (39.20 – 40.20 Mb) explained the largest proportion of the total σ_a^2 of PL with 2.96%, and eleven genes were annotated in this region (Table 4). The second most relevant genomic window was in BTA15 (73.16 – 74.16 Mb) and it explained 2.48% of σ_a^2 with nine genes located in this region (Table 4).

The C1q and tumor necrosis factor related protein 3 (*C1QTNF3*, BTA20), RAB11 family interacting protein 5 (*RAB11FIP5*, BTA11), and phosphofructokinase (*PFKM*, BTA5) genes were found in the enrichment analyses to influence the regulation of Hormone Secretion (GO:0046883) and Hormone Transport (GO:0009914) (Table 3). In the KEGG enrichment for PL, the Glucagon Signaling pathway (bta04922) was observed with the phosphofructokinase (*PFKM*, BTA5) and calmodulin 2 (*CALM2*, BTA11) genes associated with this mechanism (Table 3).

Table 4 List of the top 10 genomic windows that explained the largest proportion of the direct additive genetic variance (σ_a^2) for pregnancy loss (PL) in Nellore cattle.

20	39.20 – 40.20	*DNAJC21, BRIX1, RAD1, TTC23L, U1, RAI14, C1QTNF3, U6, AMACR, SLC45A2, ADAMTS12*	2.96
15	73.16 – 74.16	*API5, TTC17, bta-mir-670, bta-mir-6528, MIR129-2, HSD17B12, U6, ALKBH3, ACCSL*	2.48
8	57.50 – 58.50	*TLE1*	1.97
11	29.12 – 30.12	*SOCS5, MCFD2, TTC7A, STPG4, CALM2, EPCAM, MSH2, KCNK12MSH6*	1.86
5	31.40 – 32.40	*OR8S15, OR8S1, OR8S24, C5H12orf54, OR8S2, OR8S26, OR8S3, OR5BJ2, ZNF641, OR10AD1, CCDC184, ASB8, PFKM, SENP1, COL2A1, TMEM106C, VDR*	1.67
1	31.51 – 32.50*	-	1.35
23	33.52 – 34.51	*NRSN1, SNORA70, U6, FTL*	1.22
15	69.61 – 70.61	*LRRC4C*	1.10
11	11.37 – 12.36	*RAB11FIP5, SFXN5, EMX1, SPR, EXOC6B*	1.09
5	16.11 – 17.11*	-	1.06

[1]BTA: *Bos taurus* autosome; [2]All genes listed in this table were included in the enrichment analyses for this trait; *No genes were found to be in this genomic window.

2.3.3. Genome-wide association for stillbirth

The top 10 genomic windows explained 11.71% of the σ_a^2 for SB and overlap with 73 genes (63 protein-coding and 10 snRNA). The chromosomes containing these genomic regions were BTA1, BTA9, BTA11, BTA12, BTA17, BTA18, and BTA27 (Table 5 and Figure 1). The interactions between the protein-coding genes annotated for SB are presented in Supplementary File 4, Figure S3. For SB, the most relevant genomic window was in BTA1 (91.86 – 92.86), and it accounted for 2.90% of σ_a^2 (Table 5 and Figure 1). The N-acetylated alpha-linked acidic dipeptidase like 2 (*NAALADL2*, BTA1) was the only gene found in this location. The significant biological processes annotated in the GO analysis (Table 3) were Regulation of Embryonic Development (GO:0045995) with nocturnin (*NOCT*, BTA17) and delta like canonical Notch ligand 1

(*DLL1*, BTA9) genes, and Cerebellum Development (GO:0021549) with delta like canonical Notch ligand 1 (*DLL1*, BTA9) and somatostatin receptor 2 (*SSTR2*, BTA19) genes. No metabolic pathways in the KEGG enrichment were found to be associated with the genes annotated in the genomic windows associated with SB (P-value < 0.05).

Table 5 List of the top 10 genomic windows that explained the largest proportion of the direct additive genetic variance (σ_a^2) for stillbirth (SB) in Nellore cattle.

BTA[1]	Location (Mb)	Genes[2]	σ_a^2 (%)
1	91.86 – 92.86	*NAALADL2*	2.90
9	103.52 – 104.51	*DYNLT2, ERMARD, DLL1, FAM120B, PSMB1, PDCD2, OR8B1AK, OR8AR1, OR8G3M, OR10D1M, OR8B60, OR8B1Q*	1.85
1	157.51 – 158.50	*PRDM9, U1, OR2B28, ZSCAN23, TXLNA, KPNA6, TMEM39B, KHDRBS1, ASMT*	1.43
27	31.18 – 32.18	*UNC5D, KCNU1*	1.19
19	57.43 – 58.42	*SDK2, CDC42EP4, MTNAP1, VCF1, COG1, SSTR2, SLC39A11*	0.87
17	17.94 – 18.94	*MAML3, MGST2, SETD7, RAB33B, NAA15, NDUFC1, MGARP, ELF2, NOCT, U6,*	0.77
11	0.09 – 1.09	*ZC3H6, ZC3H8, FBLN7, TMEM87B, MERTK, SNORA70, ANAPC1, U6*	0.73
12	3.92 – 4.91	*U1, U6*	0.68
18	64.81 – 65.81	*U6, VN1R1, ZSCAN4, ZNF606, ZNF135, ZNF329, ZNF274, ZNF8, A1BG, RPS5, RNF225, ZNF584, ZNF132, SLC27A5, ZBTB45, TRIM28, CHMP2A, UBE2M, MZF1*	0.68
1	12.22 – 13.22*	-	0.67

[1]BTA: *Bos taurus* autosome; [2]All genes listed in this table were included in the enrichment analyses for this trait; *No genes were found to be in this genomic window.

2.3.4. Genome-wide association for pre-weaning mortality

The top 10 genomic windows that most explained the σ_a^2 and σ_m^2 of PWM were responsible for 12.03% and 12.22% of σ_a^2 and σ_m^2, respectively (Tables 6 and 7 and Figure 2). In these locations, 64 (1 MicroRNA, 58 protein-coding, 3 snoRNA, and 2 snRNA) and 49 (1 MicroRNA, 44 protein-coding, 1 snoRNA, and 3 snRNA) genes were annotated for σ_a^2 and σ_m^2, respectively. A total of 33 genes were associated

exclusively with the σ_a^2 of PWM, while 18 genes were associated solely with the σ_m^2 of PWM, and 31 genes were identified as common to both effects (Figure 3). The complete list of PWM-associated genes, indicating whether each gene was specific to direct or maternal effects, or shared by both, are presented in Supplementary File 5, Table S2. The interactions between the protein-coding genes annotated for PWM are presented in Supplementary File 6, Figure S4.

The window located in BTA 25 (7.40 – 8.40 Mb) contributed the most to the total σ_a^2 (2.65%) and σ_m^2 (2.69%) of PWM. In this genomic region, the transmembrane protein 114 (T*MEM114*, BTA25), methyltransferase 22, Kin17 lysine (*METTL22*, BTA25), 4-aminobutyrate aminotransferase (*ABAT*, BTA25), transmembrane protein 186 (*TMEM186*, BTA25), phosphomannomutase 2 (*PMM2*, BTA25), calcium regulated heat stable protein 1 (*CARHSP1*, BTA25), LITAF domain containing (*LITAFD*, BTA25), ubiquitin specific peptidase 7 (*USP7*, BTA25), RNA, U6 small nuclear 1 (*U6*, BTA25), and HUWE1 associated protein modifying stress responses (*HAPSTR1*, BTA) genes were annotated.

The significant biological processes associated with PWM (Table 3) were Regulation of Glucose Metabolic Process (GO:0010906) with the forkhead box O1 (*FOXO1*, BTA12), parkin RBR E3 ubiquitin protein ligase (*PRKN*, BTA9), and ubiquitin specific peptidase 7 (*USP7*, BTA25) genes associated with this GO term, Zinc Ion Homeostasis (GO:0055069) with the parkin RBR E3 ubiquitin protein ligase (*PRKN*, BTA9) and ATPase copper transporting beta (*ATP7B*, BTA12) genes linked to this process, Lactation (GO:0007595) with the NME/NM23 nucleoside diphosphate kinase (*NME1*, BTA19) and ATPase copper transporting beta (*ATP7B*, BTA12) genes annotated, and finally, the Regulation of Insulin Secretion (GO:0050796) with the 4-aminobutyrate aminotransferase (*ABAT*, BTA25), forkhead box O1 (*FOXO1*, BTA12), and parkin RBR E3 ubiquitin protein ligase (*PRKN*, BTA9) genes associated with this biological process. No metabolic pathways in the KEGG enrichment were found to be associated with the genes annotated in the genomic windows that most influenced PWM.

Table 6 List of the top 10 genomic windows that explained the largest proportion of the direct additive genetic variance (σ_a^2) for pre-weaning mortality (PWM) in Nellore cattle.

BTA[1]	Location (Mb)	Genes[2]	σ_a^2 (%)
25	7.40 – 8.40	*TMEM114, METTL22, ABAT, TMEM186, PMM2, CARHSP1, LITAFD, USP7, U6, HAPSTR1*	2.65
12	21.06 – 22.05	*U6, WDFY2, DHRS12, CCDC70, ATP7B, ALG11, NEK5, NEK3, CKAP2, VPS36, THSD1, U8, SLC25A15, MRPS31, SNORA71, FOXO1*	1.64
8	111.46 – 112.45	*MYT1L, PXDN, TPO, SNTG2*	1.32
12	0.51 – 1.51*	-	1.16
27	35.08 – 36.07	*IDO2, TCIM, bta-mir-2284q, ZMAT4*	1.05
3	96.33 – 97.33*	-	1.01
9	97.53 – 98.53	*PRKN, PACRG*	0.86
1	150.53 – 151.53	*KCNJ15, ERG, SNORA70, ETS2, PIK3R4*	0.79
19	57.63 – 58.63	*SDK2, CDC42EP4, MTNAP1, VCF1, COG1, SSTR2, SLC39A11*	0.78
19	35.28 – 36.28	*UTP18, MBTD1, NME2, NME1, SPAG9, TOB1, WFIKKN2, LUC7L3, ANKRD40, ABCC3, CACNA1G, SPATA20, EPN3, MYCBPAP, RSAD1, CHAD, ACSF2*	0.77

[1]BTA: *Bos taurus* autosome; [2]All genes listed in this table were included in the enrichment analyses for this trait; *No genes were found to be in this genomic window.

Table 7 List of the top 10 windows that explained the largest proportion of the maternal additive genetic variance (σ_m^2) for pre-weaning mortality (PWM) in Nellore cattle.

BTA[1]	Location (Mb)	Genes[2]	σ_m^2 (%)
25	7.40 – 8.40	*TMEM114, METTL22, ABAT, TMEM186, PMM2, CARHSP1, LITAFD, USP7, U6, HAPSTR1*	2.69
12	20.54 – 21.54	*U6, FAM124A, SERPINE3, INTS6, WDFY2, DHRS12, CCDC70, ATP7B, ALG11, NEK5*	1.66
8	111.46 – 112.45	*MYT1L, PXDN, TPO, SNTG2*	1.32
12	0.51 – 1.51*	-	1.25
27	35.08 – 36.07	*IDO2, TCIM, bta-mir-2284q, ZMAT4*	1.02
3	118.55 – 119.55	*NDUFA10, OR6B2, OR6B3, OR9S41, OR9S3, OR9S29, OR9S33, OR9S43B, OR9S32, OR9S39, OR9S44P, OR9S36B, CSF2RA, IL3RA, P2RY8*	0.95
27	44.59 – 45.58	*U6*	0.93
3	96.33 – 97.33*	-	0.83
9	97.53 – 98.53	*PRKN, PACRG*	0.79
1	150.53 – 151.53	*KCNJ15, ERG, SNORA70, ETS2, PIK3R4*	0.78

[1]BTA: *Bos taurus* autosome; [2]All genes listed in this table were included in the enrichment analyses for this trait.*No genes were found to be in this genomic window.

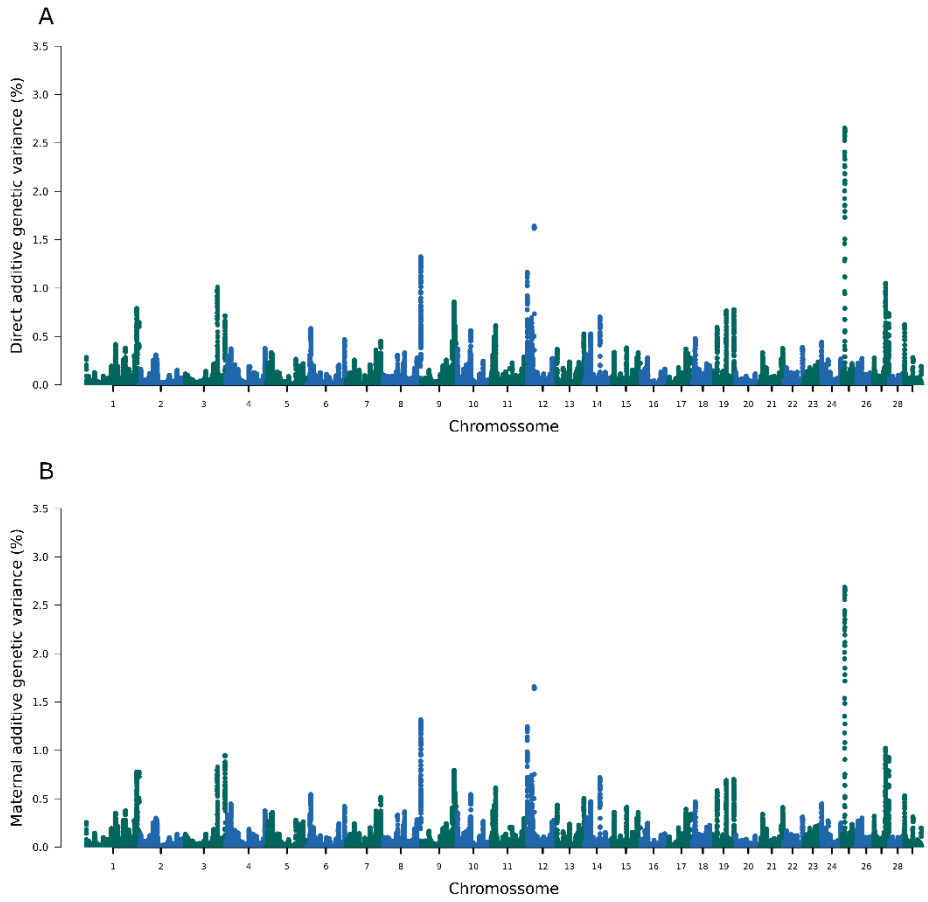

Figure 2 Manhattan plots for the proportion of the total direct (A) and maternal (B) additive genetic variance explained by 1 Mb genomic windows for pre-weaning calf mortality in Nellore cattle.

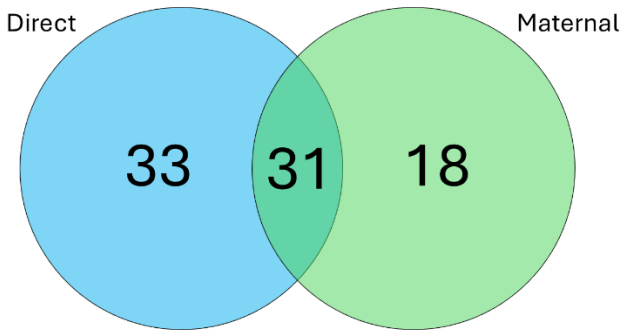

Figure 3 Venn diagram illustrating shared and specific genes between direct and maternal additive genetic effects underlying pre-weaning calf mortality in Nellore cattle.

2.4. Discussion

There are few GWAS reports for CS, PL, SB, and PWM in the Nellore breed (*Bos taurus indicus*). This lack of research may be attributed to the non-inclusion of these traits in Nellore commercial breeding programs, as their complexity makes them difficult to measure in commercial farms. Despite the low heritability estimates of these traits [33, 34], substantial genetic progress can still be obtained with genomic selection based on large reference populations. In this study, direct heritability estimates of 0.07 ± 0.01 (CS), 0.06 ± 0.03 (PL), 0.15 ± 0.03 (SB), and 0.11 ± 0.03 (PWM) were observed (Table 1).

Understanding the genetic background of these reproductive traits is of great relevance for more accurate breeding using genomic information. The identification of genetic markers associated with economically important traits can enable the refinement of genomic selection models including pre-selected SNPs [35–37]. This approach provides greater accuracy in estimated breeding values, which is particularly important for traits with low heritability, those expressed later in the animals' lives, or sex specific [35]. Hence, the present study identified genomic regions, biological processes, and metabolic pathways associated with complex reproductive traits, whose genetic background has not yet been fully elucidated in Nellore cattle.

Considering the GWAS results for CS, PL, SB, and PWM, most of the genes found are classified as protein-coding genes. Protein-coding genes are segments of DNA that are transcribed into mRNA and then translated into proteins, which are the functional molecules responsible for various biological processes [38]. In cattle, these genes play important roles in physiological pathways that regulate reproductive traits such as fertility, embryo development, and hormone signaling [39]. Mutations in protein-coding genes can influence reproduction by altering the structure or expression of the resulting proteins, potentially leading to differences in reproductive performance [38].

2.4.1. Genome-wide association for conception success

The top 10 genomic windows that most explained 17.03% of the total σ_a^2 of CS overlapping 79 genes. This finding reinforces that, while CS is greatly influenced by the environmental conditions in which animals are raised, it is also regulated by multiple genomic regions that contribute to the σ_a^2 underlying them. The genomic

window on BTA7 (3.03 - 4.03 Mb) explained 3.37% of σ_a^2. Twenty genes were found in this region (Table 2), including Wnt family member 9A (*WNT9A*), Lysophosphatidic acid receptor 2 (*LPAR2*), Zinc finger protein 354B (*ZNF354B*), and GATA zinc finger domain containing 2A (*GATAD2A*). The *WNT9A* gene plays a critical role in the WNT signaling pathway, which is essential for cell proliferation, differentiation, and embryonic development, particularly in reproductive tissues [40]. *LPAR2* is involved in lysophosphatidic acid signaling, which influences embryo implantation and placental development [41, 42]. *ZNF354B* and *GATAD2A* are involved in DNA binding and chromatin remodeling, processes that regulate gene expression important for reproductive functions [43, 44]. The identification of these genes in this genomic window underscores their potential role in contributing to the genetic background underlying CS in Nellore cattle.

The significant biological processes annotated in the functional analyses for CS were Somitogenesis (genes *PRKDC* and *WNT3A*), Somite Development (genes *PRKDC* and *WNT3A*), and Chromosome Segregation (genes *MAU2*, *BIRC5*, *MIS18A*, and *SYCE1*). Somitogenesis is a process that starts during early embryo development in vertebrates in which somites, segmented structures in the developing embryo, form and differentiate into skeletal muscle, vertebrae, and dermis [45, 46]. This mechanism is crucial for proper embryonic development, and disruptions in this activity can affect the overall health and viability of the embryo, thereby influencing CS success and maintenance [45, 46].

Some studies in humans have reported that the expression of the *WNT3A* gene at early embryogenesis has a significant influence on somitogenesis during somite development, and consequently, on the establishment of the embryonic tail [47]. Issues during this process can lead to early embryo loss, and influence CS [47]. The *PRKDC* gene was associated with structural chromosome abnormality in humans, which is an important cause of recurrent conception failure [48]. Taken together, this information along with the findings of the present study, somitogenesis and the *WNT3A* and *PRKDC* genes may play a role in regulating CS in cattle, possibly during the early stages of embryonic growth.

Chromosome segregation is a crucial biological process in the development of gametes [49]. Anomalies in chromosome segregation during the first meiotic division can hinder embryo implantation in the uterine cavity or cause spontaneous abortion [50, 51]. In cattle, disturbances in chromosome segregation have been reported to

contribute to the formation of genetically unbalanced gametes, leading to increased early embryo loss due to reduced viability [51, 52]. Therefore, this mechanism may be associated with CS due to its direct connection with reproductive mechanisms.

The expression of the *BIRC5* gene is high in tissues undergoing rapid cell proliferation during fetal development, and studies in cattle have reported its association with germinal differentiation in antral follicles [53]. In dairy cattle, the *MIS18A* gene was found in a haplotype region associated with prenatal death [54]. In mice, a homozygous deletion of the *MIS18A* gene was linked to embryo lethality [55]. For the *SYCE1* gene, a mutation in its transcription in mice was correlated with ovarian insufficiency, causing infertility in carrier females [56]. All these reports on the functions of these genes suggest that they may be associated with CS and embryonic viability, which could explain the findings of the analyses conducted in the present study for CS.

2.4.2. Genome-wide association for pregnancy loss

Regarding PL, 57 genes were annotated considering the genomic locations of the top 10 windows that most explained this trait σ^2_a. These windows were presented in 7 chromosomes and contributed to 17.76% of PL σ^2_a. The *DNAJC21, BRIX1, RAD1, TTC23L, U1, RAI14, C1QTNF3, U6, AMACR, SLC45A2, ADAMTS12* genes were found in the genomic window that most explained the σ^2_a of PL (BTA 20, 39.20 – 40.20 Mb). Among these genes, *RAI14* was identified via GWAS for semen quality traits in the Assaf sheep breed [57]. This gene may be associated with PL, as poor semen quality can contribute to improper embryonic formation during the early stages of development [58]. The *ADAMTS12* gene has been reported as a potential candidate for spontaneous preterm birth in humans [59], and its expression is associated with morphogenesis during embryonic development [60]. There are few to no reports on the association of the other genes in this genomic window with reproductive traits, providing limited evidence of their potential involvement in PL.

The biological processes identified in the gene set enrichment analysis for PL included Regulation of Hormone Secretion (GO:0046883) and Hormone Transport (GO:0009914). These processes were regulated by the *C1QTNF3*, *RAB11FIP5*, and *PFKM* genes. Hormone secretion and hormone transport can influence PL due to their roles in progesterone production during gestation, as this hormone is essential for

maintaining pregnancy [61–63]. Studies in both dairy and beef cattle have revealed an association between pregnancy maintenance, calving, and the progesterone concentrations observed during gestation [64, 65]. This suggests that issues with progesterone secretion and transport may contribute to PL in Nellore cattle, which supports the findings of this study.

The *C1QTNF3* gene has been reported to influence ovarian functions during folliculogenesis and hormone concentrations in mice [66]. Some studies also indicated that this gene plays a significant role in the bovine endometrium during early pregnancy regulating progesterone levels [67]. The *RAB11FIP5* gene was identified in a GWAS for reproductive traits in Nellore cattle, specifically associated with age at first calving and heifer early calving by 30 months [68]. This suggests that *RAB11FIP5* may influence reproductive traits in cattle, particularly in maintaining pregnancy, as age at first calving and heifer early calving are traits dependent on a full-term gestation. Regarding the *PFKM* gene, this gene has been shown to promote the proliferation and migration of trophoblast cells and the secretion of hormones essential for fetal implantation in pigs [69]. Embryo implantation is a critical stage of gestation, determining whether the pregnancy will continue, and hormone production is a pivotal aspect during gestation [69]. Considering these aspects, the findings of the present study suggest that *PFKM* may also be a major gene associated with PL in Nellore cattle, due to its role in regulating biological processes such as hormone production during pregnancy.

The metabolic pathway identified in the KEGG enrichment for PL was the Glucagon Signaling pathway (bta04922) with the genes *PFKM* and *CALM2* annotated to be responsible for regulating this process. Glucagon signaling is fundamental for glucose homeostasis, as increased plasma glucagon levels lead to elevated hepatic glucose production [70]. In humans, a lack of glucagon signaling during pregnancy has been correlated with severe intrauterine growth restriction and increased prenatal mortality [71]. During gestation, nutrients are transferred to the fetus via the placenta. However, if glucose levels are not balanced, energy transfer to the embryo may be compromised, leading to inefficient fetal growth, which can result in lethality and pregnancy loss [71–73].

The *PFKM* gene has also been found to regulate the Glucagon Signaling pathway. It plays a key role in regulating glycolysis in humans and is crucial for the expression of several enzymes involved in glucose metabolism [74, 75]. Additionally,

CALM2 is closely associated with muscle glycogen phosphorylase, which catalyzes the first step of glycogenolysis to meet energy requirements for muscle activity and is linked to glycogen storage disorders [76]. Together, these findings provide valuable insights into the genetic mechanisms underlying PL in Nellore cattle, emphasizing the importance of hormone regulation and energy metabolism in maintaining gestation.

2.4.3. Genome-wide association for stillbirth

The GWAS for SB revealed relevant genomic windows located across seven chromosomes. These regions overlapped with 73 genes and accounted for 11.71% of the total σ_a^2. This finding underscores the polygenic inheritance pattern of SB, suggesting that multiple genes with small effects collectively regulate the biological processes associated with SB occurrence.

The most relevant genomic window for SB was located on BTA1 (91.86 – 92.86 Mb) and accounted for 2.90% of the total σ_a^2. The *NAALADL2* gene was annotated in this region, which is a protein-coding gene predicted to be involved in proteolysis [77]. In Holstein cattle, this gene has been significantly associated with mean corpuscular hemoglobin [78]. Notably, this hematological parameter has been used as a predictor of SB occurrence in sows during late pregnancy [79]. This suggests that *NAALADL2* may have an indirect effect on SB occurrence by influencing corpuscular hemoglobin levels, thus reinforcing the findings of this study.

The biological processes associated with SB that were annotated include the Regulation of Embryonic Development (GO:0045995; *NOCT* and *DLL1* genes) and Cerebellum Development (GO:0021549; *DLL1* and *SSTR2* genes). The regulation of embryonic development is a significant mechanism that occurs during pregnancy and has been associated with the stillbirth phenomenon. Abnormal fetal growth, placental abruption, and Mendelian diseases are some of the processes that can disrupt embryonic development, potentially leading to SB [80–82]. Additionally, certain mutations in cattle have been reported to affect embryonic development, resulting in PL and SB by compromising the embryo's ability to properly develop until birth [83]. The *DLL1* gene, which was involved in the regulation of embryonic development, has been shown to regulate endothelial identity in mouse fetal arteries [84]. A mutation in this gene negatively impacted blood circulation, compromising fetal development [84].

The cerebellum development may be a biological process associated with SB, as cerebral malformations have been linked to SB due to the disruption of vital processes essential for the animal's development during the final stages of pregnancy [62, 85]. The cerebellum development of the fetus during gestation can be influenced by genetic diseases, metabolic disorders, infections, and mutations [86]. The *SSTR2* gene, identified in the enrichment analyses as influencing cerebellum development was also found in a GWAS for pre-weaning mortality in Nellore cattle [87]. This supports the findings of the present study, suggesting that cerebellum development during gestation may play a role in the occurrence of SB in Nellore cattle.

2.4.4. Genome-wide association for pre-weaning mortality

The most relevant genomic windows associated with the direct and maternal effects of PWM overlapped with 64 and 49 genes, respectively. This result underscores that the genetic variation of this trait is influenced by many genes with small effects, highlighting its polygenic inheritance. The T*MEM114*, *METTL22*, *ABAT*, *TMEM186*, *PMM2*, *CARHSP1*, *LITAFD*, *USP7*, *U6*, and *HAPSTR1* genes were in the genomic region that most contributed to the σ^2_a and σ^2_m of PWM (BTA25, 7.40 - 8.40 Mb). *TMEM114, TMEM186*, and *PMM2* are associated with protein glycosylation and energy metabolism, critical for early growth and development [88–90]. Using RNA sequencing, the expression of the *TMEM186* gene was correlated with brain damage in mice during the neonatal period, which contributed to increased mortality in the pre-weaning phase [91]. The *METTL22* gene is linked to methylation processes that regulate gene expression and protein function, which can also affect growth and development [92].

We observed that 31 genes were associated with both the σ^2_a and σ^2_m of PWM. This result can be attributed to several factors, including the genetic covariance between the direct and maternal additive genetic effects (σ_{am}) and the pleiotropic nature of some genes, which may influence multiple traits or biological processes linked with the expression of these genetic components underlying PWM. However, 33 genes were associated exclusively with the σ^2_a, while 18 genes were associated exclusively with the σ^2_m of PWM. The *FOXO1* gene was annotated exclusively for the σ^2_a, and it plays a pivotal role in regulating various metabolic pathways, including those involved in adipose tissue formation and skeletal muscle growth [93, 94]. Its

involvement in these processes may establish a connection to PWM by influencing energy metabolism and growth rates in cattle, which could, in turn, affect neonatal survival. The role of *FOXO1* in regulating growth and carcass traits has been reported in several livestock species, where it mediates critical pathways such as PI3K-AKT-mTOR, which are essential for tissue development and overall growth regulation [93–95]. Regarding the genes associated only with σ^2_m on PWM, one particularly relevant gene is *CSF2RA*, which encodes the receptor for granulocyte-macrophage colony-stimulating factor (GM-CSF). Disruption of the *CSF2RA* gene in bovine embryos has been shown to affect embryonic gene expression, influencing the likelihood of the embryo developing into a blastocyst [96]. This supports the hypothesis that *CSF2RA* plays a role in regulating embryonic development within the uterus, suggesting a possible involvement in the maternal genetic component associated with PWM.

The Regulation of Glucose Metabolic Process (GO:0010906) and Regulation of Insulin Secretion (GO:0050796) were among the biological processes identified in the enrichment analysis of genes associated with PWM. These mechanisms have already been associated with factors that may influence the occurrence of mortality during the pre-weaning period in economically important livestock species. In swine, blood glucose levels in newborn piglets have been linked to intrauterine growth restriction and pre-weaning mortality [97]. Furthermore, some studies have reported that insulin is a key component of cattle colostrum [98]. Insulin concentrations during the pre-weaning period in calves was also correlated with the development of the gastrointestinal tract, production of digestive enzymes, nutrient absorption capacity, growth rates, and pre-weaning survival [98–100].

FOXO1 and *PRKN* were among the genes identified as regulators of glucose metabolic processes and insulin secretion. *FOXO1* is a transcription factor involved in the regulation of genes associated with gluconeogenesis and glycogenolysis, which are key pathways for maintaining blood glucose levels during fasting [101]. *PRKN* is primarily known for its role in mitochondrial quality control through the process of mitophagy, but it also influences insulin secretion by maintaining mitochondrial health in β-cells, which is crucial for energy production and insulin release [102]. The *FOXO1* and *PRKN* genes have been reported to play significant roles in insulin secretion and signaling in response to dietary restriction in Holstein Friesian bulls, potentially influencing growth, hormone production, and overall performance [103]. During the pre-weaning of calves, issues in the glucose homeostasis mechanism can lead to

energy deficiencies and increased susceptibility to infections or other stressors, contributing to higher mortality rates [104, 105].

Zinc ion homeostasis was also identified as a biological process associated with PWM, with the *PRKN* and *ATP7B* genes playing significant roles in its regulation. In ruminants, zinc bioavailability is essential for digestion, absorption, and nutrient metabolism within the gastrointestinal tract [106]. Disruptions in zinc homeostasis can impair body development and lead to metabolic disorders [106]. Such disturbances may increase the risk of pre-weaning mortality in cattle by compromising growth, immune function, and overall health during early life stages.

PRKN influences zinc transport and distribution within cells, a vital aspect of cellular function, particularly in tissues that require high metabolic activity [102]. Disruptions in *PRKN* function can lead to imbalances in intracellular zinc levels, impairing critical metabolic processes, including energy production and immune response [102]. In pre-weaning calves, such imbalances could result in inadequate growth, increased susceptibility to infections, and higher mortality rates, underscoring the importance of proper zinc ion regulation during early development. The *PRKN* gene was found through a GWAS to influence muscle growth and fat metabolism in Guilan cattle [107].

ATP7B is another gene involved in metal ion homeostasis, specifically regulating the balance of copper and zinc ions within the body [108]. Mutations or dysfunctions in *ATP7B* can lead to toxic accumulations or deficiencies of zinc ions, disrupting metabolic processes and causing oxidative stress [108, 109]. The functions of this gene have been observed to be differentially expressed in cattle consuming a copper-deficient diet, influencing zinc metabolism and contributing to lower performance and growth [110]. Lastly, lactation was also identified as a biological process in the network of genes associated with PWM. The cow's ability to produce milk and care for her calves has previously been linked to PWM in beef cattle [111]. Cows under nutritional restriction, with mammary gland issues, or significantly affected by negative energy balance after calving tend to have lighter calves at weaning and a higher likelihood of pre-weaning mortality, as milk can be the primary source of nutrition for calves during this period [112, 113].

The *NME1* and *ATP7B* genes were observed to influence the lactation process. The *NME1* gene has been reported to be associated with milk production and mastitis in dairy cattle studies [114, 115]. Mastitis is a disease with a significant impact on both

the quantity and quality of bovine maternal milk produced, which can affect aspects of calf development during the pre-weaning period [116]. The *ATP7B* gene has been linked with feed efficiency in Holstein calves during the pre-weaning phase [117], which can be indirectly associated with PWM, as nutrient utilization through feeding is essential for the development and viability of calves until weaning [118].

2.4.5. Study implications, limitations, and prospective directions

In this study, we aimed to explore the biological aspects of CS, PL, SB, and PWM in Nellore cattle, enhancing our understanding of the genetic basis of these reproductive traits. By identifying key genomic regions and candidate genes, we are laying the groundwork for more precise breeding strategies to improve reproductive efficiency. Integrating these findings into breeding programs could boost productivity, reduce economic losses, and enhance animal welfare.

One limitation of the present study is the specific nature of the population sampled, which consisted of a Nellore cattle population from an experimental breeding program. While this design allows for precise data collection and evaluation of complex reproductive traits, the experimental setting may limit the generalizability of the results to larger commercial herds or other breeds of beef cattle. Another aspect of this study is the absence of functional validation of the candidate genes identified through the GWAS. Although significant genomic windows and candidate genes associated with reproductive traits were highlighted, the lack of direct experimental validation prevents confirmation of causal relationships. Future research incorporating transcriptomic, proteomic, or gene-editing approaches would be valuable to determine the functional roles of these genes in reproductive processes.

While our study employed statistical methods to control for multiple testing, such as the FDR procedure to minimize the risk of false positives during the GO and KEGG enrichment, we acknowledge that significant processes involving a small number of genes should be interpreted with caution. The biological relevance of such findings may be uncertain, particularly in categories with a limited number of genes (e.g., two genes), where statistical power may be low. In this context, future studies should aim to replicate our findings in larger, independent, and more diverse populations to strengthen the validity of the identified associations.

Looking forward, functional validation of the identified genes should be a primary focus for future research. This could be achieved through gene expression analyses in reproductive tissues or via gene-editing technologies to test the phenotypic effects of specific genetic variants. Studies in diverse environmental contexts would also provide knowledge into gene-environment interactions that may modulate the expression of reproductive traits.

2.5. Conclusions

Important genomic regions associated with conception success, pregnancy loss, stillbirth, and pre-weaning mortality were identified in the present study, providing a deeper biological understanding of these reproductive traits in Nellore cattle. In general, the genes annotated within the most relevant genomic windows were involved in biological processes and metabolic pathways related to somitogenesis, embryo development, hormone secretion, insulin and glucose regulation, zinc ion homeostasis, and lactation. The identification of these genes and mechanisms is of paramount importance for the search for causal mutations that have a significant influence on the expression of reproductive traits and can contribute to breeding and future genomic selection.

2.6. Supplementary files

All supplementary materials have been deposited in the Harvard Dataverse repository and are available at the permanent link: https://doi.org/10.7910/DVN/T9HCLU

2.7. Declarations

2.7.1. Ethics approval and consent to participate

Animal Care and Use Committee approval was not obtained for this study because all the analyses were performed using pre-existing datasets.

2.7.2. Consent for publication

Not applicable

2.7.3. Availability of data and materials

The data can be accessed by contacting MEZM (e-mail: mezmercadante@gmail.com) upon a reasonable request for research purposes.

2.7.4. Competing interests

The authors declare that they have no competing interests.

2.7.5. Funding

São Paulo Research Foundation (FAPESP) for financial support (2017/10630-2 and 2017/50339-5) and providing grants to GRDR (2023/11176-4 and 2024/01909-7), LFMM (2022/11852-7), and JPSV (2024/05697-4).
Coordination for the Improvement of Higher Education Personnel (CAPES, Finance Code 001) for providing grant to MSB.
National Council for Science and Technological Development (CNPq) for providing grants to LEF, LGA and MEZM.

2.7.6. Authors' contributions

MEZM coordinated the team and supervised all the stages of the study. GRDR, JNSGC, and MEZM conceived and designed the study. GRDR, LFMM, and MEZM conducted the data analyses. LEF, LGA and MEZM obtained the resources for this research. GRDR, LFMM, JNSGC, JPSV, LFB, JBSN, MSB, FMM, LEF, LGA, LFB, and MEZM contributed to the data acquisition and interpretation of the results. GRDR, LFMM, JNSGC, LFB, LFB, and MEZM wrote and edited the manuscript. All authors reviewed and contributed to the editing of the manuscript and approved its final version.

2.7.7. Acknowledgements

The authors thank the staff of the Institute of Animal Science (Sertãozinho, SP, Brazil) for their efforts in collecting the data used in this study.

References

1. De Vries A. Reproduction: Economics of reproductive efficiency. Rev Colomb Cienc Pecu. 2017;30:218–21.

2. Lancaster PA, Larson RL. Evaluation of Strategies to Improve the Environmental and Economic Sustainability of Cow–Calf Production Systems. Animals. 2022;12:1–15.

3. Amaya A, Martínez R, Cerón-Muñoz M. Selection indexes using principal component analysis for reproductive, beef and milk traits in Simmental cattle. Trop Anim Health Prod. 2021;53:1–6.

4. Berry DP, Wall E, Pryce JE. Genetics and genomics of reproductive performance in dairy and beef cattle. Animal. 2014;8:105–21.

5. Costa E V., Ventura HT, Veroneze R, Silva FF, Pereira MA, Lopes PS. Estimated genetic associations among reproductive traits in Nellore cattle using Bayesian analysis. Anim Reprod Sci. 2020;214:1–6.

6. Ogunbawo AR, Mulim HA, Campos GS, Oliveira HR. Genetic Foundations of Nellore Traits: A Gene Prioritization and Functional Analyses of Genome-Wide Association Study Results. Genes. 2024;15:1131.

7. Zhang H, Wang Z, Wang S, Li H. Progress of genome wide association study in domestic animals. J Anim Sci Biotechnol. 2012;3:1–10.

8. Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, et al. Genome-wide association studies. Nat Rev Methods Primers. 2021;1:1–21.

9. Benfica LF, Brito LF, do Bem RD, de Oliveira LF, Mulim HA, Braga LG, et al. Detection and characterization of copy number variation in three differentially-selected Nellore cattle populations. Front Genet. 2024;15:1–12.

10. Bem RD, Benfica LF, Silva DA, Carrara ER, Brito LF, Mulim HA, et al. Assessing different metrics of pedigree and genomic inbreeding and inbreeding effect on growth, fertility, and feed efficiency traits in a closed-herd Nellore cattle population. BMC Genomics. 2024;25:1–14.

11. Rodrigues GRD, Rezende VT, Mercadante MEZ, Bonilha SFM, Canesin RC, Raineri C, et al. Animal growth models as a tool to estimate resilience indicators in *Bos indicus* and *Bos taurus* heifers: Selection effects and genetics parameters. Livest Sci. 2024;282:1–9.

12. Mota LFM, Santos SWB, Júnior GAF, Bresolin T, Mercadante MEZ, Silva JA V, et al. Meta-analysis across Nellore cattle populations identifies common metabolic mechanisms that regulate feed efficiency-related traits. BMC Genomics. 2022;23:424.

13. Benfica LF, Brito LF, do Bem RD, Mulim HA, Glessner J, Braga LG, et al. Genome-wide association study between copy number variation and feeding behavior, feed efficiency, and growth traits in Nellore cattle. BMC Genomics. 2024;25.

14. Sargolzaei M, Chesnais JP, Schenkel FS. A new approach for efficient genotype imputation using information from relatives. BMC Genomics. 2014;15:1–12.

15. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience. 2020;9:1–9.

16. Mota LFM, Arikawa LM, Santos SWB, Fernandes Júnior GA, Alves AAC, Rosa GJM, et al. Benchmarking machine learning and parametric methods for genomic prediction of feed efficiency-related traits in Nellore cattle. Sci Rep. 2024;14.

17. Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee DH. BLUPF90 AND RELATED PROGRAMS (BGF90). In: 7th World Congress on Genetics Applied to Livestock Production. 2022. p. 1–2.

18. Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J Dairy Sci. 2010;93:743–52.

19. VanRaden PM. Efficient Methods to Compute Genomic Predictions. J Dairy Sci. 2008;91:4414–23.

20. Geweke J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. Staff Report. 1991;148:1–8.

21. Smith BJ. boa: An R Package for MCMC Output Convergence Assessment and Posterior Inference. J Stat Softw. 2007;21:1–37.

22. Wang H, Misztal I, Aguilar I, Legarra A, Muir WM. Genome-wide association mapping including phenotypes from relatives without genotypes. Genet Res (Camb). 2012;94:73–83.

23. Yin L, Zhang H, Tang Z, Xu J, Yin D, Zhang Z, et al. rMVP: A Memory-efficient, Visualization-enhanced, and Parallel-accelerated Tool for Genome-wide Association Study. Genomics Proteomics Bioinformatics. 2021;19:619–28.

24. da Silva Neto JB, Peripoli E, Pereira ASC, Stafuzza NB, Lôbo RB, Fukumasu H, et al. Weighted genomic prediction for growth and carcass-related traits in Nelore cattle. Anim Genet. 2023;54:271–83.

25. Carvalho FE, Espigolan R, Berton MP, Neto JBS, Silva RP, Grigoletto L, et al. Genome-wide association study and predictive ability for growth traits in Nellore cattle. Livest Sci. 2020;231.

26. Irano N, De Camargo GMF, Costa RB, Terakado APN, Magalhães AFB, Silva RMDO, et al. Genome-wide association study for indicator traits of sexual precocity in Nellore cattle. PLoS One. 2016;11.

27. Mota LFM, Carvajal AB, Bernardes PA, Buzanskas ME, Baldi F, Lôbo RB, et al. Integrating genome-wide association study and pathway analysis reveals physiological aspects affecting heifer early calving defined at different ages in Nelore cattle. Genomics. 2022;114:110395.

28. Magalhã AFB, De Camargo GMF, Junior Fernandes GA, Gordo DGM, Tonussi RL, Costa RB, et al. Genome-Wide Association Study of Meat Quality Traits in Nellore Cattle. PLoS One. 2016;11.

29. Arikawa LM, Mota LFM, Schmidt PI, Frezarim GB, Fonseca LFS, Magalhães AFB, et al. Genome-wide scans identify biological and metabolic pathways regulating carcass and meat quality traits in beef cattle. Meat Sci. 2024;209.

30. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation. 2021;2:1–9.

31. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple. Journal of The Royal Statistical Society. 1995;57:289–300.

32. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: Protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43:447–52.

33. Schmidek A, José M, Paranhos Da Costa R, Eugênia M, Mercadante Z, Macedo De Toledo L, et al. Genetic and non-genetic effects on calf vigor at birth and preweaning mortality in Nellore calves. Revista Brasileira de Zootecnia. 2013;42:421–7.

34. Silva LCM, Baldi F, Aboujaoude C, Venturini GC, Albuquerque LG, Paranhos da Costa MJR. Genetic parameter estimates for prenatal and postnatal mortality in Nellore cattle. Journal of Animal Breeding and Genetics. 2017;134:27–33.

35. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME. Invited review: Genomic selection in dairy cattle: Progress and challenges. Journal of Dairy Science. 2009;92:433–43.

36. Montaldo HH, Casas E, Ferraz JBS, Vega-Murillo VE, Román-Ponce SI. Opportunities and challenges from the use of genomic selection for beef cattle breeding in Latin America. Animal Frontiers. 2012;2:23–9.

37. Mancin E, Mota LFM, Tuliozi B, Verdiglione R, Mantovani R, Sartori C. Improvement of genomic predictions in small breeds by construction of genomic relationship matrix through variable selection. Front Genet. 2022;13 May:1–25.

38. Lee TI, Young RA. TRANSCRIPTION OF EUKARYOTIC PROTEIN-CODING GENES. 2024.

39. Ortega MS. Identification of genes associated with reproductive function in dairy cattle. Anim Reprod. 2018;15:923–32.

40. Xiang Y, Lin G, Zhang Q, Tan Y, Lu G. Knocking down Wnt9a mRNA levels increases cellular proliferation. Mol Biol Rep. 2008;35:73–9.

41. Ye X, Chun J. Lysophosphatidic acid (LPA) signaling in vertebrate reproduction. Trends in Endocrinology and Metabolism. 2010;21:17–24.

42. Ye X, Herr DR, Diao H, Rivera R, Chun J. Unique uterine localization and regulation may differentiate LPA3 from other lysophospholipid receptors for its role in embryo implantation. Fertil Steril. 2011;95.

43. Torchy MP, Hamiche A, Klaholz BP. Structure and function insights into the NuRD chromatin remodeling complex. Cellular and Molecular Life Sciences. 2015;72:2491–507.

44. Sharifi Tabar M, Giardina C, Feng Y, Francis H, Moghaddas Sani H, Low JKK, et al. Unique protein interaction networks define the chromatin remodelling module of the NuRD complex. FEBS Journal. 2022;289:199–214.

45. Maroto M, Bone RA, Kim Dale J. Somitogenesis. Development. 2012;139:2453–6.

46. Wang Y, Cai H, Luo X, Ai Y, Jiang M, Wen Y. Insight into unique somitogenesis of yak (Bos grunniens) with one additional thoracic vertebra. BMC Genomics. 2020;21:1–9.

47. Yamanaka Y, Hamidi S, Yoshioka-Kobayashi K, Munira S, Sunadome K, Zhang Y, et al. Reconstituting human somitogenesis in vitro. Nature. 2023;614:509–20.

48. Cheng Z, Cheng D, Li J, Guo L, Zhang W, Zhang C, et al. Polymorphisms Within DNA Double-Strand Breaks Repair-Related Genes Contribute to Structural Chromosome Abnormality in Recurrent Pregnancy Loss. Front Genet. 2021;12:1–11.

49. Greaney J, Wei Z, Homer H. Regulation of chromosome segregation in oocytes and the cellular basis for female meiotic errors. Hum Reprod Update. 2018;24:135–61.

50. Wasielak-Politowska M, Kordowitzki P. Chromosome Segregation in the Oocyte: What Goes Wrong during Aging. Int J Mol Sci. 2022;23:1–8.

51. Jennings RL, Griffin DK, O'Connor RE. A new approach for accurate detection of chromosome rearrangements that affect fertility in cattle. Animals. 2020;10:1–13.

52. Raudsepp T, Chowdhary BP. Chromosome aberrations and fertility disorders in domestic animals. Annu Rev Anim Biosci. 2016;4:15–43.

53. Sirard MA. Somatic environment and germinal differentiation in antral follicle: The effect of FSH withdrawal and basal LH on oocyte competence acquisition in cattle. Theriogenology. 2016;86:54–61.

54. Fritz S, Capitan A, Djari A, Rodriguez SC, Barbat A, Baur A, et al. Detection of Haplotypes Associated with Prenatal Death in Dairy Cattle and Identification of Deleterious Mutations in GART, SHBG and SLC37A2. PLoS One. 2013;8:1–8.

55. Kim IS, Lee M, Park KC, Jeon Y, Park JH, Hwang EJ, et al. Roles of Mis18α in Epigenetic Regulation of Centromeric Chromatin and CENP-A Loading. Mol Cell. 2012;46:260–73.

56. Hernández-López D, Geisinger A, Trovero MF, Santiñaque FF, Brauer M, Folle GA, et al. Familial primary ovarian insufficiency associated with an SYCE1 point mutation: Defective meiosis elucidated in humanized mice. Mol Hum Reprod. 2020;26:485–97.

57. Serrano M, Ramón M, Calvo JH, Jiménez M, Freire F, Vázquez JM, et al. Genome-wide association studies for sperm traits in Assaf sheep breed. Animal. 2021;15.

58. Franco GA, Peres RFG, Martins CFG, Reese ST, Vasconcelos JLM, Pohler KG. Sire contribution to pregnancy loss and pregnancy-associated glycoprotein production in nelore cows. J Anim Sci. 2018;96:632–40.

59. Mani S, Ghosh J, Lan Y, Senapati S, Ord T, Sapienza C, et al. Epigenetic changes in preterm birth placenta suggest a role for ADAMTS genes in spontaneous preterm birth. Hum Mol Genet. 2019;28:84–95.

60. Oztas E, Ozler S, Ersoy AO, Erkenekli K, Sucak A, Ergin M, et al. Placental ADAMTS-12 Levels in the Pathogenesis of Preeclampsia and Intrahepatic Cholestasis of Pregnancy. Reproductive Sciences. 2016;23:475–81.

61. Diskin MG, Waters SM, Parr MH, Kenny DA. Pregnancy losses in cattle: Potential for improvement. Reprod Fertil Dev. 2016;28:83–93.

62. Sigdel A, Bisinotto RS, Peñagaricano F. Genes and pathways associated with pregnancy loss in dairy cattle. Sci Rep. 2021;11:13329.

63. Smith BD, Poliakiwski B, Polanco O, Singleton S, De Melo GD, Muntari M, et al. Decisive points for pregnancy losses in beef cattle. Reproduction, Fertility and Development. 2022.

64. López-Gatius F, Santolaria P, Yániz JL, Hunter RHF. Progesterone supplementation during the early fetal period reduces pregnancy loss in high-yielding dairy cattle. Theriogenology. 2004;62:1529–35.

65. Reese ST, Franco GA, Poole RK, Hood R, Fernadez Montero L, Oliveira Filho R V., et al. Pregnancy loss in beef cattle: A meta-analysis. Anim Reprod Sci. 2020;212.

66. Mao Z, Yang L, Lu X, Tan A, Wang Y, Ding F, et al. C1QTNF3 in the murine ovary and its function in folliculogenesis. Reproduction. 2018;155:333–46.

67. Musavi SAA, Yamashita S, Fujihara T, Masaka H, Islam MR, Kim S, et al. Analysis of differentially expressed genes and the promoters in bovine endometrium throughout estrus cycle and early pregnancy. Animal Science Journal. 2018;89:1609–21.

68. Sbardella AP, Watanabe RN, Costa RM da, Bernardes PA, Braga LG, Rey FSB, et al. Genome-wide association study provides insights into important genes for reproductive traits in nelore cattle. Animals. 2021;11.

69. Wang Y, Zhou C, Meng F, Hu Q, Ding Y, Wang X, et al. Ssc-miR-92b-3p Regulates Porcine Trophoblast Cell Proliferation and Migration via the PFKM Gene. Int J Mol Sci. 2022;23.

70. Janah L, Kjeldsen S, Galsgaard KD, Winther-Sørensen M, Stojanovska E, Pedersen J, et al. Glucagon receptor signaling and glucagon resistance. International Journal of Molecular Sciences. 2019;20.

71. Ouhilal S, Vuguin P, Cui L, Du X-Q, Gelling RW, Reznik SE, et al. Hypoglycemia, hyperglucagonemia, and fetoplacental defects in glucagon receptor knockout mice: a role for glucagon action in pregnancy maintenance. Am J Physiol Endocrinol Metab. 2012;302:522–31.

72. Qiao L, Saget S, Lu C, Zang T, Dzyuba B, Hay WW, et al. The Essential Role of Pancreatic a-Cells in Maternal Metabolic Adaptation to Pregnancy. Diabetes. 2022;71:978–88.

73. Dumolt JH, Rosario FJ, Kramer AC, Horwitz S, Powell TL, Jansson T. Maternal glucagon-like peptide-1 is positively associated with fetal growth in pregnancies complicated with obesity. Clin Sci. 2023;137:663–78.

74. Quesada I, Tudurí E, Ripoll C, Nadal Á. Physiology of the pancreatic α-cell and glucagon secretion: Role in glucose homeostasis and diabetes. Journal of Endocrinology. 2008;199:5–19.

75. Zhang Y, Qin C, Wang J, Yang L, Yan X, Zhi S, et al. Phosphofructokinase family genes in grass carp: Molecular identification and tissue-specific expression in response to glucose, insulin and glucagon. Comp Biochem Physiol B Biochem Mol Biol. 2024;269.

76. Migocka-Patrzałek M, Elias M. Muscle glycogen phosphorylase and its functional partners in health and disease. Cells. 2021;10.

77. Whitaker HC, Shiong LL, Kay JD, Grönberg H, Warren AY, Seipel A, et al. N-acetyl-L-aspartyl-L-glutamate peptidase-like 2 is overexpressed in cancer and promotes a pro-migratory and pro-metastatic phenotype. Oncogene. 2014;33:5274–87.

78. Yang T, Luo H, Lou W, Chang Y, Brito LF, Zhang H, et al. Genetic background of hematological parameters in Holstein cattle based on genome-wide association and RNA sequencing analyses. J Dairy Sci. 2024;107:4772–92.

79. Hill S V, Del M, Amezcua R, Ribeiro ES, O'sullivan TL, Friendship RM. Article. 2024.

80. Diskin MG, Parr MH, Morris DG. Embryo death in cattle: An update. Reproduction, Fertility and Development. 2012;24:244–51.

81. Stanley KE, Giordano J, Thorsten V, Buchovecky C, Thomas A, Ganapathi M, et al. Causal Genetic Variants in Stillbirth. New England Journal of Medicine. 2020;383:1107–16.

82. Kajdy A, Modzelewski J, Cymbaluk-Płoska A, Kwiatkowska E, Bednarek-Jędrzejek M, Borowski D, et al. Molecular pathways of cellular senescence and placental aging in late fetal growth restriction and stillbirth. International Journal of Molecular Sciences. 2021;22.

83. Flisikowski K, Venhoranta H, Nowacka-Woszuk J, Mckay SD, Flyckt A, Taponen J, et al. A novel mutation in the maternally imprinted PEG3 domain results in a loss of MIMT1 expression and causes abortions and stillbirths in cattle (Bos taurus). PLoS One. 2010;5.

84. Sörensen I, Adams RH, Gossler A. DLL1-mediated Notch activation regulates endothelial identity in mouse fetal arteries. Blood. 2009;113:5680–8.

85. Venhoranta H, Bauersachs S, Taponen J, Lohi H, Taira T, Andersson M, et al. Fetal growth restriction caused by MIMT1 deletion alters brain transcriptome in cattle. International Journal of Developmental Neuroscience. 2013;31:463–7.

86. Grafe MR, Kinney HC. Neuropathology Associated With Stillbirth. Semin Perinatol. 2002;26:83–8.

87. Marín-Garzón NA, Magalhães AFB, Schmidt PI, Serna M, Fonseca LFS, Salatta BM, et al. Genome-wide scan reveals genomic regions and candidate genes underlying direct and maternal effects of preweaning calf mortality in Nellore cattle. Genomics. 2021;113:1386–95.

88. Cochran SD, Cole JB, Null DJ, Hansen PJ. Single nucleotide polymorphisms in candidate genes associated with fertilizing ability of sperm and subsequent embryonic development in cattle. Biol Reprod. 2013;89.

89. Yang S, Li X, Liu X, Ding X, Xin X, Congfei J, et al. Parallel comparative proteomics and phosphoproteomics reveal that cattle myostatin regulates phosphorylation of key enzymes in glycogen metabolism and glycolysis pathway. 2018.

90. Mei C, Wang H, Liao Q, Khan R, Raza SHA, Zhao C, et al. Genome-wide analysis reveals the effects of artificial selection on production and meat quality traits in Qinchuan cattle. Genomics. 2019;111:1201–8.

91. Zhou H, Wang X, Cheng R, Hou X, Chen Y, Feng Y, et al. Analysis of long non-coding RNA expression profiles in neonatal rats with hypoxic-ischemic brain damage. J Neurochem. 2019;149:346–61.

92. Cloutier P, Lavallée-Adam M, Faubert D, Blanchette M, Coulombe B. Methylation of the DNA/RNA-binding protein Kin17 by METTL22 affects its association with chromatin. J Proteomics. 2014;100:115–24.

93. Kousteni S. FoxO1, the transcriptional chief of staff of energy metabolism. Bone. 2012;50:437–43.

94. Ioannilli L, Ciccarone F, Ciriolo MR. Adipose tissue and FOXO1: Bridging physiology and mechanisms. Cells. 2020;9.

95. Kanlisi RA, Amuzu-Aweh EN, Naazie A, Otsyina HR, Kelly TR, Gallardo RA, et al. Genetic architecture of body weight, carcass, and internal organs traits of Ghanaian local chickens. Front Genet. 2024;15:1–20.

96. Sosa F, Uh K, Drum JN, Stoecklein KS, Davenport KM, Ortega MS, et al. Disruption of CSF2RA in the bovine preimplantation embryo reduces development and affects embryonic gene expression in utero. Reproduction and Fertility. 2023;4:1–10.

97. Staarvik T, Framstad T, Heggelund M, Brynjulvsrud Fremgaarden S, Kielland C. Blood-glucose levels in newborn piglets and the associations between blood-glucose levels, intrauterine growth restriction and pre-weaning mortality. Porcine Health Manag. 2019;5:1–10.

98. McGrath JJ. Accelerated pre-weaning growth rates in dairy calves: do antioxidants have a place? Anim Prod Sci. 2016;56:1275–84.

99. Swali A, Cheng Z, Bourne N, Wathes DC. Metabolic traits affecting growth rates of pre-pubertal calves and their relationship with subsequent survival. Domest Anim Endocrinol. 2008;35:300–13.

100. Wu S, Li X, Chen X, Zhu Y, Yao J. Optimizing the growth and immune system of dairy calves by subdividing the pre-weaning period and providing different milk volumes for each stage. Animal Nutrition. 2021;7:1296–302.

101. Kousteni S. FoxO1, the transcriptional chief of staff of energy metabolism. Bone. 2012;50:437–43.

102. Terešak P, Lapao A, Subic N, Boya P, Elazar Z, Simonsen A. Regulation of PRKN-independent mitophagy. Autophagy. 2022;18:24–39.

103. Keogh K, Kenny DA, Kelly AK, Waters SM. Insulin secretion and signaling in response to dietary restriction and subsequent re-alimentation in cattle. Physiol Genomics. 2015;47:344–54.

104. Trefz FM, Feist M, Lorenz I. Hypoglycaemia in hospitalised neonatal calves: Prevalence, associated conditions and impact on prognosis. Veterinary Journal. 2016;217:103–8.

105. Saraiva DT, Moreira SS, Santos MEP, Almeida ER, Rennó LN, Valadares Filho S de C, et al. Performance and Metabolic Responses of Nellore Cows Subjected to Different Supplementation Plans during Prepartum. Animals. 2024;14:2283–93.

106. Mir SH, Mani V, Pal RP, Malik TA, Sharma H. Zinc in Ruminants: Metabolism and Homeostasis. Proceedings of the National Academy of Sciences India Section B - Biological Sciences. 2020;90:9–19.

107. Jourshari MG, Shadparvar AA, Hossein-Zadeh NG, Rafeie F, Banabazi MH, Johansson AM. Genome-wide association study on abdomen depth, head width, hip width, and withers height in native cattle of Guilan (Bos indicus). PLoS One. 2023;18:1–8.

108. DiDonato M, Narindrasorasak S, Forbes JR, Cox DW, Sarkar B. Expression, purification, and metal binding properties of the N-terminal domain from the Wilson disease putative copper-transporting ATPase (ATP7B). Journal of Biological Chemistry. 1997;272:33279–82.

109. Meacham KA, Cortés MP, Wiggins EM, Maass A, Latorre M, Ralle M, et al. Altered zinc balance in the Atp7b−/− mouse reveals a mechanism of copper toxicity in Wilson disease. Metallomics. 2018;10:1595–606.

110. Dermauw V, De Cuyper A, Duchateau L, Waseyehon A, Dierenfeld E, Clauss M, et al. A disparate trace element metabolism in zebu (Bos indicus) and crossbred

(Bos indicus × Bos taurus) cattle in response to a copper-deficient diet 1. J Anim Sci. 2014;92:3007–17.

111. Iewdiukow M, Lema OM, Velazco JI, Quintans G. Is it possible to accurately estimate lactation curve parameters in extensive beef production systems? Applied Animal Science. 2020;36:509–14.

112. Axford MM, Khansefid M, Haile-Mariam M, Goddard ME, Pryce JE. Genetic evaluation for stillbirth and pre-weaning mortality in Australian dairy cattle. J Dairy Sci. 2024;107:1–15.

113. Barry J, Bokkers EAM, de Boer IJM, Kennedy E. Pre-weaning management of calves on commercial dairy farms and its influence on calf welfare and mortality. Animal. 2020;14:2580–7.

114. Ogorevc J, Kunej T, Razpet A, Dovc P. Database of cattle candidate genes and genetic markers for milk production and mastitis. Anim Genet. 2009;40:832–51.

115. Ma Y, Khan MZ, Xiao J, Alugongo GM, Chen X, Chen T, et al. Genetic markers associated with milk production traits in dairy cattle. Agriculture. 2021;11:1–25.

116. Royster E, Wagner S. Treatment of Mastitis in Cattle. Veterinary Clinics of North America - Food Animal Practice. 2015;31:17–46.

117. Salleh SM, Mazzoni G, Løvendahl P, Kadarmideen HN. Gene co-expression networks from RNA sequencing of dairy cattle identifies genes and pathways affecting feed efficiency. BMC Bioinformatics. 2018;19:1–10.

118. Wang C, Li D, Yang J, Xia Y, Tu Y, White R, et al. Weaning performance of beef cattle calves based on concentrate intake. Animals. 2020;10:1–18.

CHAPTER 3: EFFECT OF GENOMIC REGIONS HARBORING PUTATIVE LETHAL HAPLOTYPES ON REPRODUCTIVE PERFORMANCE IN CLOSED EXPERIMENTAL SELECTION LINES OF NELLORE CATTLE

Gustavo R. D. Rodrigues[1,2,3*], Joslaine N. S. G. Cyrillo[2], Lúcio F. M. Mota[1], Patrícia I. Schmidt[1], Júlia P. S. Valente[1,2], Eduarda S. Oliveira[1], Lúcia G. Albuquerque[1,4], Luiz F. Brito[3], Maria E. Z. Mercadante[1,2,4*]

[1]São Paulo State University, School of Agriculture and Veterinary Science, Jaboticabal, SP 14884-900, Brazil
[2]Institute of Animal Science, Beef Cattle Research Center, Sertãozinho, SP 14174-000, Brazil
[3]Department of Animal Sciences, Purdue University, West Lafayette, IN 47907, USA
[4]National Council for Science and Technological Development, Brasilia, DF 71605-001, Brazil

***Corresponding authors**: Gustavo R. D. Rodrigues: grd.rodrigues@unesp.br and Maria E. Z. Mercadante: mezmercadante@gmail.com

Published in Scientific Reports (ISSN: 2045-2322)

Received: 05 August 2024
Accepted: 28 January 2025
Published: 03 February 2025

Cite this article: Rodrigues, G.R.D., Cyrillo, J.N.S.G., Mota, L.F.M. *et al.* Effect of genomic regions harboring putative lethal haplotypes on reproductive performance in closed experimental selection lines of Nellore cattle. *Sci Rep* **15**, 4113 (2025). https://doi.org/10.1038/s41598-025-88501-7

Abstract

Lethal alleles are mutations in the genome that cause embryonic losses in affected homozygous embryos and, therefore, can negatively influence reproduction rates in commercial populations. Thus, this study aimed to identify genomic regions containing potential lethal haplotypes in Nellore breed; identify candidate genes located within these regions; and investigate the reproductive performance of heterozygous carriers of lethal haplotypes in Nellore cattle. Forty-five genomic regions harboring putative lethal haplotypes were identified, which overlap with 360 genes. Gene ontology analyses of these genes revealed biological processes associated with the development of sexual traits in males and females, key functions of the immune system, energy homeostasis, and embryonic development. The gene networks were involved in metabolic pathways including ovarian steroidogenesis, oocyte meiosis, and insulin secretion. Matings between carrier dam and carrier sire led to a reduction of up to -203.46% in pregnancy success probability, an increase of 275.15% in probability of pregnancy loss, 295.03% for stillbirth occurrence, and 301.40% for pre-weaning mortality when compared to non-carrier dam and sire matings. The results highlight the importance of identifying animals that are carriers of lethal haplotypes to avoid the propagation of these haplotypes in the population.

Keywords: pregnancy success, recessive alleles, stillbirth, Zebu cattle

3.1. Background

Most mutations in diploid organisms typically manifest as recessive alleles, resulting in a harmful effect when homozygous [1]. Lethal alleles have been identified to significantly impact the profitability and efficiency of livestock industry. These genetic variants have the potential to reduce the expression of economically important traits [2], cause embryonic losses [3], reduce pregnancy and rebreeding rates [4, 5], and increase post-natal mortality [5].

Lethal alleles can be tracked using a haplotype-based approach, where the harmful region in the genome can be inferred based on observed and expected frequencies of homozygous haplotypes in the population [6]. This approach uses data from living animals, instead of relying on information from affected embryos (that may not have survived). Its efficacy depends on having a sufficient number of genotyped animals in a population, as well as accurate genealogical records for tracing inheritance patterns [5–7]. If the number of genotyped animals is sufficiently high, the complete absence of homozygous carriers of some haplotypes in different segments of the genome may not be a completely random event [6].

In populations under direct selection for economically important traits, the presence of a lethal allele conferring an advantage to heterozygous individuals in their phenotypic performance, may facilitate the propagation of this deleterious genetic variant across subsequent generations [2, 7]. This event hides the antagonistic effects associated with the homozygous recessive state of the allele, which results in harmful effects or mortality among affected animals [6]. Furthermore, higher linkage disequilibrium between the lethal allele and alleles with favorable effects on the target traits in breeding programs may also maintain the lethal allele in the population due to selection practices [2].

Most studies aimed at identifying lethal genetic variants in bovine species have been focused on taurine breeds (*Bos taurus taurus*), with few reports in animals of Zebu origin (*Bos taurus indicus*). Schmidt et al. [5] reported 30 potentially lethal haplotypes in the Nellore breed (*Bos taurus indicus*), distinct from those identified in taurine breeds. Similarly, Id-Lahoucine et al. [4] identified 19 potentially lethal haplotypes in the Angus breed based on a transmission ratio distortion (TRD) approach and observed that heterozygous carriers of these haplotypes showed a reduction of 15% in pregnancy rates. These results support the hypothesis that

genomic regions harboring putative lethal haplotypes may have undesirable effects even in heterozygous carriers. Furthermore, identifying carriers of lethal haplotypes to guide mating decisions may be an effective strategy to prevent the spread of lethal alleles to future generations, potentially enhancing reproductive efficiency.

In this study, we used data belonging to a closed experimental Nellore breeding program selecting animals for stabilizing and directional selection, based on YW for 44 years [8, 9]. These closed selection lines differ from each other on phenotypic performance, such as average body weight at different ages, body measurements and carcass trait. This distinctive breeding design can offer valuable insights into the implications of the harmful effect of recessive alleles on reproductive aspects within a closed Nellore population under selection. Hence, the objectives of this study were to identify potential genomic regions containing lethal haplotypes in a Nellore *(Bos taurus indicus)* cattle population; to find genes and functional processes linked to the gene network within specified genomic regions harboring putative candidates for lethal haplotypes; and to investigate the reproductive performance of carriers of lethal haplotypes, with a specific focus on the probabilities of pregnancy success, pregnancy losses, stillbirth, and pre-weaning mortality in at-risk matings.

3.2. Materials and Methods

3.2.1. Ethical statement

The production system evaluated followed all animal welfare guidelines established by Law No. 11.977 of the State of São Paulo, Brazil. Data were obtained from an existing database, and therefore, approval from the Ethics Committee was not required.

3.2.2. Population

The Nellore dataset belonged to an experimental selection design from the Beef Cattle Research Center (Institute of Animal Science – IZ, Sertãozinho, SP, Brazil; Figure 4). In 1980, the Nellore IZ experimental breeding program was established with the aim of selecting animals based on YW, measured at 378 days of age in young bulls and 550 days of age in heifers. Three distinct selection lines were designed: Line 1 - Nellore control (NeC), with animals selected for YW close to the average of the

contemporary group (stablishing selection), within birth year x herd; Line 2 - Nellore Selection (NeS), with animals selected for the YW with a maximum selection differential, within birth year x herd; and Line 3 - Nellore Traditional (NeT) with animals selected for higher YW, and since 2008 has been also selected for lower residual feed intake (RFI), within birth year x herd.

The selection lines are closed, which means that only bulls and dams born in the respective selection line are used for breeding. The inbreeding levels of all three herds are controlled with planned mattings considering the co-ancestry coefficient. The bulls are utilized in mating seasons for two consecutive years, beginning at two years of age with 15 cows and at three years of age with 30 cows. Annually, on average, 4, 6, and 8 bulls are utilized for breeding in the NeC, NeS, and NeT selection lines, respectively.

Figure 4 Animals of the Nellore breed from the experimental breeding program of the Institute of Animal Science - Beef Cattle Research Center (Sertãozinho, SP, Brazil)

3.2.3. Pedigree and genomic databases

The pedigree dataset included genealogical information of 12,843 individuals born from 1950 to 2022, involving 478 sires and 2927 dams. We used the RelaX2 software [10] for pedigree evaluation and it was identified a total of 13 founders for NeC, 27 for NeS, and 29 for NeT. The estimated effective population size was 107 ± 15 animals, considering the three selection lines, based on the methodology outlined by Gutiérrez et al. [11]. Additionally, the number of equivalent complete generations (ECG) was 5.02, while the number of discrete generations (DGE) was 6.50 for the three selection lines. The mean inbreeding was 2.81% for NeC, 2.23% for NeS, and 2.15% for NeT, with the highest inbred animal exhibiting 25.00% of inbreeding.

The genomic dataset included 3,226 genotyped animals, with 1,741 males and 1,485 females genotyped with different Bead chip assays (Supplementary File 1, Table S1). Before imputation, were removed monomorphic markers, markers with the same coordinate, located on non-autosomal chromosomes and with GenCall score lower than 0.90. After this quality control (QC), the animals genotyped with medium-density panels (50K and 75K SNP panels) were imputed to the Illumina BovineHD panel (770K) using the FImpute v.3 software [12]. The expected accuracy from the imputation process was higher than 0.97 [13, 14]. After the imputation, 3,226 animals and 612,154 SNP markers remained in the genomic dataset for further analyses. Possible genotype inconsistencies between parents and progeny were adjusted using the seekparentf90 software [15].

3.2.4. Reproductive records

A total of 5,093 reproductive records (pregnancy success, pregnancy loss, stillbirth and pre-weaning mortality) from 1,258 cows and 190 sires genotyped were used for the analyses. Annually, cows and heifers were exposed to natural or artificial insemination at a fixed time during 90-day breeding season. After approximately 30 days from the end of the breeding season, an ultrasound evaluation was performed to diagnose pregnancy, and females were classified as pregnant (1, success) or non-pregnant (0, failure). The pregnancy loss was determined by attributing a value of 1 to females with a positive pregnancy diagnosis but did not calve and 0 otherwise. For non-pregnant cows, no information about pregnancy loss was given.

The stillbirths were considered as calves that died within the first 48 hours after birth being assigned as 1 (stillbirth), and calves that survived beyond this period were assigned as 0 (non-stillborn). Calves were weaned at around 7 months and pre-weaning mortality was determined by assigning a value of 1 (mortality) for calves that did not reach weaning age and 0 (survived) for animals that survived until weaning.

Contemporary groups (CG) for pregnancy success, pregnancy loss, and stillbirth were based on the combination of the breeding year (1980 to 2022) and selection line (NeC, NeS, and NeT). We did not include the season information in CG because all cows were bred in a single season of the year. For pre-weaning mortality, CG consisted of calves' birth year (1981 to 2023), sex (male or female), and selection line (NeC, NeS, and NeT). CGs with less than three animals and no phenotypic variability within CG were excluded from further analyses.

3.2.5. Identification of candidates for lethal haplotypes

To identify potential candidates for lethal haplotypes this study followed the methods proposed by VanRaden et al. [6] and applied by Schmidt et al. [5]. The findhap.f90 software v.3 [16] was used to obtain the haplotypes using the sliding windows method [17, 18]. The haplotype construction involved three-step iterative processes: initially, haplotypes consisting of 2,000 markers on the same chromosome were identified; next, haplotypes comprising 632 markers; and, finally, it identified haplotypes containing up to 200 markers, which were subsequently used and considered for further analyses. These numbers were defined based on the recommendations of findhap.f90 software considering the SNP panel density used for this study [16].

The haplotypes that showed frequencies higher than 2% in the population but were not observed in a homozygous state were selected for further evaluation as putative lethal haplotypes. We established this threshold value to select only the haplotypes present in a significant number of genotyped animals. This was done with the aim of identifying genetic variants with the highest lethal potential within the population. This threshold value was also applied by Hoff et al. [7] and Schmidt et al. [2].

In accordance with the methods described by VanRaden et al. [6], two approaches were used to estimate the expected number of homozygous individuals

for each haplotype: 1) the Simple method – assumes random mating across the population over time. The calculation involves dividing the number of genotyped individuals by 4 and then multiplying it by the square of the carrier frequency of the haplotypes; and 2) the Mating method – accounts for the actual mating patterns observed in the population that generated the genotyped individuals. In this approach, the expected number of homozygous individuals is calculated by dividing the number of carrier mating sire × carrier maternal grandsire pairs by 4. This approach assumes that the allele frequencies for maternal granddams and maternal grandsires are equal.

The probabilities of observing zero homozygotes when n > 0 is expected were obtained using analogous formulas applied by VanRaden et al. [6], Jenko et al. [2], and Schmidt et al. [5]. For the Simple method, this probability (Phh) was calculated as $(1 – C^2/4) \times N$, where the absence of homozygous animals depends on the carrier frequency of heterozygous animals (C) and the number of genotyped animals (N). In contrast, for the Mating method, the probability follows a Bernoulli distribution and is equal to 0.75 raised to the power of the observed number of carrier mating sire × carrier maternal grandsire pairs.

To identify a putative recessive lethal haplotype region among the thousands of haplotypes that could present zero homozygous individuals by chance, specific conditions were set following Wu et al. [3]: 1) the haplotype carrier frequency had to exceed 2%; 2) the number of expected homozygous individuals for the haplotype had to be greater than 1 and 3) the probabilities of observing zero homozygotes had to be less than 0.6. All haplotypes satisfying these conditions were selected for further analyses.

3.2.6. Functional and gene set enrichment

The genomic regions identified as potential hotspots for lethal haplotype candidates were mapped to the genes located within these regions using the BioMart tool from the ENSEMBL (https://www.ensembl.org/biomart/martview/), considering the ARS-UCD 1.2 *Bos taurus* assembly [19]. Functional classification of genes for biological mechanisms (Gene Ontology - GO) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways were identified using the cluster- Profiler R Package [20].

3.2.7. Statistical analyses

After identifying all potential lethal haplotypes and their respective carriers, four mating classes were defined according to the carrier status of each lethal haplotype: 1) noncarriers ($\mathrm{NC_{Dam}} \times \mathrm{NC_{Sire}}$); 2) carrier dam with noncarrier sire ($\mathrm{C_{Dam}} \times \mathrm{NC_{Sire}}$); 3) noncarrier dam with carrier sire ($\mathrm{NC_{Dam}} \times \mathrm{C_{Sire}}$); and 4) carrier dam with carrier sire ($\mathrm{C_{Dam}} \times \mathrm{C_{Sire}}$). Following these mating classes, a logistic regression model considering a logit link function, was used to describe the effects of the four mating classes on the probabilities of pregnancy success, pregnancy loss, stillbirth, and pre-weaning mortality:

$$\boldsymbol{\eta} = \mathbf{X_m m} + \mathbf{X_{CG} CG} + \mathbf{X_{ca} ca} + \mathbf{e}$$

where $\boldsymbol{\eta}$ was the logit transformation of the phenotypic records (θ) with $\eta_i = \ln\left(\frac{\theta_i}{1-\theta_i}\right)$, assumed as $\mathbf{y} \sim bin(1, \theta)$; $\mathbf{y}$ is the vector of phenotypic records (pregnancy, pregnancy loss, stillbirth, and pre-weaning mortality), in which *i*th element in $\mathbf{y}$ followed a Bernoulli distribution $y_i \sim Be(\theta_i)$; $\mathbf{m}$ is a vector with the fixed effects of mating class; $\mathbf{X_m}$ is an incidence matrix relating phenotypic records to mating class effect for each lethal haplotype; $\mathbf{CG}$ is a vector of fixed effects of CG; $\mathbf{X_{CG}}$ is an incidence matrix relating phenotypic records to the CG effect; $\mathbf{ca}$ is a vector with the linear and quadratic effects of cow's age at breeding as a covariate; $\mathbf{X_{ca}}$ is an incidence matrix relating phenotypic records to the cow's age effects; $\mathbf{e}$ is the residual vector assumed as $\mathbf{e} \sim N(0, \sigma_e^2)$, where σ_e^2 is the variance of the residuals.

The effects of mating type were expressed as an odds ratio probability, comparing the elements m_1 with the other elements (m_2, m_3, and m_4) of $\mathbf{m}$ corresponding mating classes. All the analyses described in this section were performed using Rstudio version 4.3.2 with the package stats [21].

3.3. Results

3.3.1. Candidates for lethal haplotypes

A total of 45 genomic regions located on 18 chromosomes exhibited lethal haplotypes with an observed frequency higher than 2% and with no observed homozygous haplotypes in the genotyped animals (Table 8).

The BTAs that exhibited at least 2 potential candidate lethal haplotypes were identified on: BTA1 (7 candidates), BTA2 (4 candidates), BTA7 (3 candidates), BTA10

(2 candidates), BTA13 (4 candidates), BTA15 (3 candidates), BTA17 (4 candidates), BTA18 (4 candidates), BTA26 (3 candidates), BTA27 (2 candidates), and BTA29 (2 candidates). A total of 45 lethal candidate haplotypes presented probabilities lower than 0.60 using both, the simple and mating detection methods. Among these 45 putative lethal haplotypes, 42 had not been reported in other studies with cattle as potential carriers of lethal alleles. Only the haplotypes 998.14 in BTA7 (segment 998, haplotype 14), 2213.8 in BTA17 (segment 2213, haplotype 8), and 2620.22 in BTA 22 (segment 2620, haplotype 22) were previously identified as lethal haplotypes [5].

The lethal haplotype with the highest number of carriers (735 animals) was the 1947.12 haplotype in BTA15 (segment 1947, haplotype 12), with size of 0.74 Mb and a frequency of 11.39% in the population. For this haplotype, 26 homozygous individuals were expected considering the Simple method (random mating) and 43 homozygous individuals considering the Mating method (observed mating pattern). The lethal haplotype with the lowest number of carriers was located in BTA26 (2850.17; segment 2850, haplotype 17), which has a size of 1.06 Mb, 224 carriers and a frequency of 3.48% in the evaluated population.

The lethal haplotypes exhibited an average size of 0.91 ± 0.43 Mb, with the largest haplotype identified on BTA10 (1387.10; segment 1387, haplotype 10), with a size of 3.42 Mb, 367 carriers and a frequency of 5.69% in genotyped the population. The smallest was haplotype 1637.10 located in BTA12 (segment 1637, haplotype 10), with a size of 0.57 Mb, 245 carriers and a frequency of 3.80% in the population.

Table 8 List of potential genomic regions candidates for lethal haplotypes in Nellore cattle.

Lethal Haplotype[1]	BTA[2]	Start location[3]	End location[3]	Size (Mb)	Carriers	Haplotype frequency (%)	Expected hh (Simple)[4]	Phh (Simple)[5]	Expected hh (Mating)[6]	Phh (mating)[7]
10.14	1	8,387,543	9,417,014	1.0295	236	3.66	3	3.39E-01	4	7.97E-02
41.4	1	35,837,523	36,849,734	1.0122	305	4.73	5	1.64E-01	7	1.46E-02
42.4	1	36,851,422	37,775,698	0.9243	310	4.81	6	1.54E-01	8	1.27E-02
43.4	1	37,780,384	38,549,159	0.7688	316	4.90	6	1.44E-01	8	1.07E-02
44.4	1	38,555,417	39,542,541	0.9871	289	4.47	4	1.99E-01	7	2.29E-02
46.4	1	40,648,407	41,576,866	0.9285	285	4.42	4	2.06E-01	6	2.49E-02
96.27	1	81,218,959	82,115,702	0.8967	278	4.30	7	2.25E-01	6	3.05E-02
193.7	2	83,872	990,397	0.9065	249	3.86	6	3.01E-01	5	6.03E-02
223.10	2	25,459,389	26,072,321	0.6129	367	5.69	9	7.34E-02	11	2.23E-03
224.9	2	26,073,608	26,830,916	0.7573	377	5.85	9	6.35E-02	11	1.59E-03
225.11	2	26,836,702	27,435,862	0.5992	370	5.74	9	7.01E-02	11	2.00E-03
935.15	6	111,034,774	111,679,609	0.6448	214	3.31	3	4.13E-01	4	1.26E-01
998.14	7	36,554,986	37,251,910	0.6969	401	6.20	10	4.50E-02	16	2.82E-03
1008.1	7	53,068,682	53,888,447	0.8198	238	3.69	4	3.33E-01	4	7.65E-02
1016.4	7	59,428,787	60,126,329	0.6975	288	4.47	5	2.00E-01	7	2.31E-02
1110.12	8	21,624,756	22,743,291	1.1185	240	3.72	3	3.27E-01	5	7.31E-02
1387.10	10	22,284,939	25,709,495	3.4246	367	5.69	8	7.35E-02	11	2.24E-03
1419.5	10	52,713,425	53,805,606	1.0922	264	4.09	4	2.59E-01	5	4.23E-02
1637.10	12	16,539,142	17,107,382	0.5682	245	3.80	3	3.12E-01	5	6.54E-02
1736.1	13	13,160,486	13,869,773	0.7093	277	4.29	5	2.27E-01	6	3.12E-02
1739.3	13	15,706,108	16,489,553	0.7834	290	4.50	3	1.95E-01	7	2.18E-02
1756.27	13	28,652,199	29,318,085	0.6659	240	3.72	7	3.27E-01	5	7.31E-02
1774.2	13	42,739,592	43,780,435	1.0408	259	4.01	8	2.73E-01	5	4.78E-02
1892.3	14	56,732,026	57,630,427	0.8984	363	5.63	15	7.74E-02	10	2.52E-03
1946.13	15	21,179,798	22,038,861	0.8591	712	11.03	24	5.38E-05	40	1.07E-10

Table 8 (Continued) List of potential genomic regions candidates for lethal haplotypes in Nellore cattle.

Lethal Haplotype[1]	BTA[2]	Start location[3]	End location[3]	Size (Mb)	Carriers	Haplotype frequency (%)	Expected hh (Simple)[4]	Phh (Simple)[5]	Expected hh (Mating)[6]	Phh (mating)[7]
1947.12	15	22,040,526	22,784,950	0.7444	735	11.39	26	2.81E-05	43	2.35E-11
1948.12	15	22,785,789	23,402,175	0.6164	712	11.04	25	5.32E-05	40	1.04E-10
2027.5	16	207,162	1,258,314	1.0512	283	4.38	3	2.12E-01	6	2.66E-02
2213.8	17	65,952,250	66,468,921	0.5167	699	10.83	17	7.72E-05	38	2.48E-10
2214.4	17	66,477,303	67,188,615	0.7113	284	4.40	3	2.10E-01	6	2.61E-02
2214.9	17	66,477,303	67,188,615	0.7113	436	6.76	6	2.49E-02	15	1.79E-04
2219.3	17	70,431,294	71,438,318	1.0070	296	4.59	7	1.83E-01	7	1.87E-02
2260.18	18	27,892,662	28,634,466	0.7418	301	4.67	3	1.73E-01	7	1.64E-02
2291.5	18	53,287,408	53,992,697	0.7053	282	4.37	4	2.15E-01	6	2.74E-02
2293.2	18	54,862,900	55,865,174	1.0023	293	4.55	3	1.89E-01	7	2.02E-02
2300.15	18	62,798,271	63,775,507	0.9772	240	3.72	5	3.27E-01	5	7.34E-02
2620.22	22	46,693,892	47,385,511	0.6916	352	5.46	3	9.05E-02	10	3.63E-03
2787.1	25	4,326,940	4,997,458	0.6705	263	4.07	4	2.62E-01	5	4.37E-02
2850.17	26	12,124,239	13,181,206	1.0570	224	3.48	3	3.77E-01	4	1.02E-01
2852.5	26	14,182,969	15,308,498	1.1255	236	3.66	3	3.39E-01	4	7.99E-02
2852.58	26	14,182,969	15,308,498	1.1255	233	3.61	3	3.49E-01	4	8.51E-02
2922.8	27	16,795,465	17,580,705	0.7852	240	3.72	3	3.27E-01	5	7.34E-02
2923.7	27	17,582,546	18,429,080	0.8465	234	3.63	3	3.45E-01	4	8.32E-02
3024.1	29	9,456,456	10,514,416	1.0580	276	4.27	5	2.30E-01	6	3.20E-02
3072.6	29	48,542,674	49,949,151	1.4065	258	4.00	3	2.76E-01	5	4.91E-02

[1]Lethal haplotypes are discerned by DNA segment number and the corresponding haplotype within segment (e.g., 2620.22 signifies segment 2620 and haplotype 22). [2]BTA = *Bos Taurus* Autosome; [3]Start and end position of the lethal haplotypes. [4]The expected number of individuals exhibiting homozygous haplotypes (*hh*) considering the simple method is equal the number of individuals genotyped (3,226) divided by four and multiplied by square of carrier frequency. [5]The probabilities of observing no haplotypes homozygous (*Phh*) is contingent upon both the carrier frequency of heterozygous individuals and the total number of genotyped animals. [6]The expected number of individuals with *hh* considering the mating method is equivalent to the number of carrier service sire × carrier maternal grandsire mattings divided by 4. [7]The *Phh* of the matting method adheres to a Bernoulli distribution and is calculated as 0.75 raised to the frequencies of the observed number of carrier service sire × carrier maternal grandsire pairs.

3.3.2. Identified genes and functional enrichment analyses

A total of 360 genes were identified within the genomic regions indicated as a putative lethal haplotype (Supplementary File 7, Table S3). Enrichment analysis identified biological processes associated with the development of sexual traits in males and females, key functions of the immune system, energy homeostasis, and embryonic development (Table 9).

Table 9 Significant Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses

Description	P-value	Gene ID
GO:0046546 - Development of primary male sexual traits	0.0097	*ZFP42, GATA3, LHB, LRP2, BAX, AKR1C3*
GO:0008584 - Male gonad development	0.0094	*ZFP42, GATA3, LHB, LRP, BAX, AKR1C3*
GO:0030540 - Female genitalia development	0.0027	*LRP2, BAX, RBP4*
GO:0042445 - Hormone metabolic process	0.0157	*DHRS9, AKR1C4, CYP26A1, GATA3, CYP26C1, LHB, RBP4, BCO2, AKR1C3*
GO:0042593 - Glucose homeostasis	0.0185	*FGF21, FOXA3, INS, TRPM4, G6PC, RBP4, TRA2B, SMARCB1, TH*
GO:0006112 - Energy reserve metabolic process	0.0134	*GYS1, IGF2, INS, PYGB, PHLDA2*
GO:0045995 - Regulation of embryonic development	0.0146	*RUVBL2, NLRP5, PHLDA2, ACTR8*
bta04913 - Ovarian steroidogenesis	0.0181	*INS, LHB, AKR1C3*
bta04114 - Oocyte meiosis	0.0307	*INS, PPP2R1B, YWHAH, CALM3*
bta04911 - Insulin secretion	0.0343	*INS, TRPM4, CACNA1D*

Specifically, the zinc finger protein (*ZFP42*), GATA binding protein 3 (*GATA3*), luteinizing hormone subunit beta (*LHB*), LDL receptor related protein 2 (*LRP2*), BCL2 associated X, apoptosis regulator (*BAX*), and aldo-keto reductase family 1 member C3 (*AKR1C3*) genes, were annotated and associated with the development of primary male sexual traits and male gonad development (GO:0046546; GO:0008584). Furthermore, the *LRP2*, *BAX*, and retinol binding protein 4 (*RBP4*) genes were identified to influence female genitalia development (GO:0030540).

Regarding the immune system, dehydrogenase/reductase 9 (*DHRS9*), aldo-keto reductase family 1 member C4 (*AKR1C4*), cytochrome P450 Family 26 subfamily A Member 1 (*CYP26A1*), *GATA3*, cytochrome P450 Family 26 subfamily C Member 1 (*CYP26C1*), *LHB*, *RBP4*, beta-carotene oxygenase 2 (*BCO2*), and *AKR1C3* genes were associated with the hormone metabolic process (GO:0042445).

Concerning energy homeostasis, the glucose homeostasis process (GO:0042593) was identified in the GO analyses, with the fibroblast growth factor 21 (*FGF21*), forkhead box A3 (*FOXA3)*, insulin (*INS*), transient receptor potential cation channel subfamily M member 4 (*TRPM4*), glucose-6-phosphatase catalytic subunit 2 (*G6PC2*), *RBP4*, transformer 2 beta homolog (*TRA2B*), SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1 (*SMARCB1*), and tyrosine hydroxylase (*TH*) genes annotated and associated with this function. The glycogen synthase 1 (*GYS1*), insulin like growth factor 2 (*IGF2*), *INS*, glycogen phosphorylase B (*PYGB*), and pleckstrin homology like domain family A member 2 (*PHLDA2*) genes influenced the energy reserve metabolic process (GO:00006112). For the regulation of embryonic development (GO:0045995) there were also identified genes such as RuvB like AAAATPase 2 (*RUVBL2*), NLR family pyrin domain containing 5 (*NLRP5*), *PHLDA2*, and actin-related protein 8 (*ACTR8*).

Gene enrichment analysis identified metabolic pathways affecting the ovarian steroidogenesis by the action of the gene set (*INS*, *LHB*, and *AKR1C3*), oocyte meiosis (*INS*, *PPP2R1B*, *YWHAH*, and *CALM3*); and insulin secretion (INS, TRPM4 and CACNA1D) (Table 9).

3.3.3. Reproductive performance of matings involving carriers of lethal haplotypes

Among the 1,258 genotyped cows with mating records, 1,209 of them were carriers of at least one candidate lethal haplotype in a heterozygous state. Regarding

the sires, 186 out of 190 were carriers of at least one putative lethal haplotype. Overall, 2,564 combinations between the mating type C_{Dam} × C_{Sire} were observed in the reproductive records (Supplementary File 8, Table S4). The number of matings for each lethal haplotype involving NC_{Dam} × NC_{Sire}, C_{Dam} × NC_{Sire}, NC_{Dam} × C_{Sire} and C_{Dam} × C_{Sire} are presented in Supplementary File 9, Table S5.

Only the haplotypes 42.4 (BTA1), 935.15 (BTA6), and 2214.4 (BTA17) exhibited associations with the four traits: pregnancy success, pregnancy loss, stillbirth and pre-weaning mortality (P-value < 0.05). On the other hand, the haplotypes 1736.1 (BTA13), 1756.27 (BTA13), 1774.2 (BTA13), 1892.3 (BTA14), 1947.12 (BTA15), 1948.12 (BTA15), 2213.8 (BTA17), 2219.3 (BTA17), 2260.18 (BTA18), 2300.15 (BTA18), 2787.1 (BTA25), 2850.17 (BTA26), and 3072.6 (BTA29) were not associated with pregnancy success, pregnancy loss, stillbirth and pre-weaning mortality occurrences (P-value > 0.05).

Twenty-one lethal haplotypes had significant effects (P-value < 0.05) on pregnancy success (Table 10). Matings involving the combination of C_{Dam} × NC_{Sire} exhibited reductions of up to -86.83% in pregnancy success probability when compared with NC_{Dam} × NC_{Sire} mating, as observed with haplotype 223.10 (BTA2). Similarly, mating NC_{Dam} × C_{Sire} showed reductions of up to -88.02% in pregnancy success when compared to NC_{Dam} × NC_{Sire} mating, detected with haplotype 193.7 (BTA2). Furthermore, mating C_{Dam} × C_{Sire} exhibited reductions of up to -203.46% in pregnancy success probability when compared to NC_{Dam} × NC_{Sire} mating considering haplotype 998.14 (BTA7).

For pregnancy loss, 11 haplotypes exhibited significant effects on its occurrence (P < 0.05) (Table 11). The matings C_{Dam} × NC_{Sire} increased the probability of pregnancy loss by 78.36% when compared to NC_{Dam} × NC_{Sire}, as observed in haplotype 2923.7 (BTA27). Similarly, mating NC_{Dam} × C_{Sire} had up to 105.23% probability of experiencing pregnancy loss, observed in carriers of haplotype 2923.7 (BTA27). Regarding matings of C_{Dam} × C_{Sire}, the probabilities of pregnancy loss occurrence ranged from 39.04%, observed in haplotype 10.14 (BTA1), to 275.15% when considering haplotype 1739.3 (BTA13).

Thirteen candidate lethal haplotypes significantly influenced stillbirth occurrence (P < 0.05), with 6 of them being located on BTA1 (Table 11). In matings between C_{Dam} × NC_{Sire} and NC_{Dam} × C_{Sire} of haplotype 2027.5 (BTA16), the probabilities of observing a stillbirth were of 112.50% and 181.13%, respectively. Furthermore, matings between

carriers of haplotype 935.15 in the $C_{Dam} \times C_{Sire}$ class showed the highest probabilities for stillbirth occurrence, with a 295.03% increase in its likelihood.

Considering pre-weaning mortality, 14 candidate lethal haplotypes had significant effects on the pre-weaning mortality of calves ($P < 0.05$) (Table 11). The probabilities of observing pre-weaning mortality ranged from 14.56% (haplotype 935.15, BTA6) to 85.43% (haplotype 2027.5, BTA16) in $C_{Dam} \times NC_{Sire}$ matings, 14.88% (haplotype 1110.12, BTA8) to 80.19% (haplotype 2852.58, BTA26) in $NC_{Dam} \times C_{Sire}$ matings, and 44.03% (haplotype 42.4, BTA1) to 301.40% (haplotype 1110.12, BTA8) in $C_{Dam} \times C_{Sire}$ matings.

Table 10 Effects[1] of mating classes considering lethal haplotypes carriers in the probabilities of pregnancy success in Nellore cattle

Lethal hap[2]	C_{Dam}×NC_{Sire}[3]	NC_{Dam}×C_{Sire}[4]	C_{Dam}×C_{Sire}[5]	P-value
10.14	-1.23	-7.45	-24.32	0.038
42.4	-4.59	-43.79	-55.14	0.020
96.27	-30.83	-35.68	-54.00	0.010
193.7	-36.59	-88.02	-106.92	0.002
223.10	-86.83	-86.00	-97.67	<0.001
935.15	-16.22	-20.83	-46.18	0.012
998.14	-6.27	-6.53	-203.46	0.018
1008.1	-21.83	-38.36	-41.46	0.007
1016.4	-22.13	-26.33	-52.24	0.025
1387.10	-7.54	-34.72	-95.25	0.018
1419.5	-40.17	-45.55	-105.79	<0.001
1637.10	-25.52	-55.35	-81.13	0.042
1739.3	-28.01	-28.48	-60.49	<0.001
1946.13	-18.40	-15.63	-32.96	0.047
2214.4	-8.26	-45.59	-56.54	0.016
2214.9	-13.88	-9.93	-57.42	0.014
2291.5	-10.52	-33.58	-66.85	<0.001
2293.2	-20.60	-37.30	-51.14	<0.001
2922.8	-21.75	-22.87	-38.38	0.013
2923.7	-28.57	-32.48	-37.94	0.009
3024.1	-34.12	-11.59	-61.78	0.008

[1]The effects of the lethal haplotypes are expressed in percentage and represents the odds ratio of the mating class compared to a scenario $NC_{dam} \times NC_{sire}$ (e.g., in lethal haplotype 10.14, the result reported in the $C_{Dam} \times C_{Sire}$ class indicates that this mating class has -24.32% chance of pregnancy success compared to the mating class $NC_{Dam} \times NC_{Sire}$); [2]Lethal haplotypes are discerned by DNA segment number and the corresponding haplotype within segment (e.g., 2620.22 signifies segment 2620 and haplotype 22); [3]$C_{Dam} \times NC_{Sire}$: carrier dam with noncarrier sire; [4]$NC_{Dam} \times C_{Sire}$: noncarrier dam with carrier sire; [5]$C_{Dam} \times C_{Sire}$: carrier dam with carrier sire.

Table 11 Effects[1] of mating classes considering lethal haplotypes carriers on the probabilities of pregnancy loss, stillbirth and pre-weaning mortality in Nellore cattle.

Lethal hap[2]	**C_{Dam}×NC_{Sire}[3]**	**NC_{Dam}×C_{Sire}[4]**	**C_{Dam}×C_{Sire}[5]**	**P-value**
Pregnancy loss				
10.14	22.41	15.73	39.04	0.04
42.4	49.96	3.9	50.91	0.023
193.7	36.95	47.61	73.66	0.018
935.15	26.87	27.15	210.81	0.02
1419.5	29.75	47.74	58.77	0.013
1739.3	40.69	53.37	275.14	0.003
2214.4	63.92	78.39	84.72	0.048
2291.5	50.35	51.15	72.63	0.002
2293.2	53.33	22.65	182.24	0.024
2922.8	71.02	103.97	161.09	0.002
2923.7	78.36	105.23	162.7	<0.001
Stillbirth				
41.4	61.22	112.4	124.42	<0.001
42.4	59.06	113.11	141.02	<0.001
43.4	58.72	95	108.47	<0.001
44.4	68.93	111.79	129.9	0.004
46.4	71.8	112.5	130.34	<0.001
193.7	26.51	43.35	187.84	0.008
935.15	50.83	51.83	295.03	<0.001
1008.1	15.84	12.14	99.12	0.045
1110.12	56.93	52.8	279.03	0.008
1739.3	15.25	16.03	89.79	0.022
2027.5	112.5	181.13	210.24	<0.001
2214.4	18.27	48.81	54.01	0.043
2620.22	86.51	109.82	144.03	<0.001
Pre-weaning mortality				
10.14	37.55	32.48	54.62	<0.001
42.4	31.79	28.06	44.03	0.003
44.4	16.77	19.39	77.83	0.03

46.4	23.74	31.38	79.84	0.041
223.1	30.03	26.36	271.72	<0.001
224.9	40.02	31.79	258.28	<0.001
225.11	38.73	25.3	268.38	<0.001
935.15	14.56	21.73	189.56	0.032
1110.12	20.33	14.88	301.4	0.04
2027.5	85.43	72.75	174.88	<0.001
2214.4	23.1	41.14	101.14	0.015
2620.22	27.14	31.48	139.84	<0.001
2852.5	23.79	27.3	49.55	0.034
2852.58	36.29	80.19	85	0.008

[1]The effects of the lethal haplotypes are expressed in percentage and represents the odds ratio of the mating class compared to a scenario $NC_{dam} \times NC_{sire}$ (e.g., in lethal haplotype 10.14 considering pregnancy loss, the result reported in the $C_{Dam} \times C_{Sire}$ class indicates that this mating class has 39.04% more chance of experiencing pregnancy loss when compared to the mating class $NC_{Dam} \times NC_{Sire}$); [2]Lethal haplotypes are discerned by DNA segment number and the corresponding haplotype within segment (e.g., 2620.22 signifies segment 2620 and haplotype 22); [3]$C_{Dam} \times NC_{Sire}$: carrier dam with noncarrier sire; [4]$NC_{Dam} \times C_{Sire}$: noncarrier dam with carrier sire; [5]$C_{Dam} \times C_{Sire}$: carrier dam with carrier sire.

3.4. Discussion

3.4.1. Candidates for lethal haplotypes

Lethal haplotypes have been identified in livestock species, including horses [22, 23], sheep [24], dairy cattle [3, 25, 26], and beef cattle [2, 5, 27]. The identification of lethal haplotypes represents a strategy for guiding mating decisions aimed at mitigating the dissemination of lethal alleles to subsequent generations and compromising herd productive efficiency. However, there is a lack of research on the Nellore breed (*Bos taurus indicus*), having only one study to date that identified potential lethal haplotypes candidates [5], despite its significant contribution to global meat production. The study of Nellore cattle is particularly relevant due to their prominent role in Brazilian beef production, as they are resilient to environmental challenges of tropical climates [9]. Furthermore, the experimental breeding program that provided the data for this study is one of the pioneering initiatives in genetic improvement of the Nellore breed in Brazil. Genetic material from elite animals is supplied to herds across the entire country, playing an important role in establishing

diverse herds throughout Brazilian territory, and some Latin American countries. This emphasizes the importance of investigating genetic factors that may influence the productivity and sustainability of Nellore cattle, such as the occurrence of lethal haplotypes.

In this study, 45 chromosomal regions containing putative lethal haplotypes were identified (Table 8). These haplotypes exhibited frequencies exceeding 2% in the population with no homozygous individuals observed among genotyped animals. Furthermore, the probabilities of observing zero homozygous individuals were lower than 0.60, as determined by the two detection tests used to search for lethal alleles (simple and mating methods) based on the expected frequency of homozygous haplotype carriers, as described by VanRaden et al. [6].

Out of the 45 candidates, 42 represent novel genomic regions that have not been previously reported in other cattle studies as hotspots for lethal alleles. This finding reinforces the hypothesis that the Nellore breed may harbor distinct genomic segments containing previously unreported deleterious mutations. Furthermore, it is important to note that the population evaluated in this study originates from a selection experiment initiated in the 1980s and has since been maintained as closed selection lines. A closed population can make high frequency haplotypes easier to find, but then they may have low frequency in the general Nellore population. Besides, populations of closed breeding programs may also encounter challenges in maintaining genetic variability across generations, leading to increased levels of inbreeding resulting from mating between genetically related animals [28]. In this study, the highest inbreed animal was observed in the selection line NeC and carried 5 potential lethal haplotypes. Increased inbreeding can lead to an increased frequency of deleterious alleles [29, 30].

The haplotypes 998.14 (BTA7), 2213.8 (BTA17), and 2620.22 (BTA22), previously identified by Schmidt et al. [5], were also found in this study as potential candidates for lethal haplotypes in the Nellore breed (Table 8). This result suggests that these haplotypes may be specific to the Nellore cattle breed and highlights potential fixation of these genomic regions in Nellore animals. Schmidt et al. [5] conducted their study using data from 276 commercial herds located in different Brazilian regions. One possible explanation for the recurrence of these three lethal haplotypes in our study is that all Nellore animals trace back to a few common

ancestors, which may contribute to the identification of these haplotypes across studies.

The genomic regions containing putative lethal haplotypes were identified on 18 chromosomes (1, 2, 6, 7, 8, 10, 12, 13, 14, 15, 16, 17, 18, 22, 25, 26, 27, and 29), with heterozygous carrier frequencies ranging from 3.48% to 11.39% in the population (Table 8). Schmidt et al. [5] reported frequencies of lethal haplotypes ranging from 2.98% to 12.21% in Nellore cattle considering 30 haplotypes harboring potential lethal alleles, while Wu et al. [26] in a Nordic Red Dairy cattle population observed frequencies ranging from 3.82% to 12.91% on 18 putative recessive lethal regions.

The haplotypes exhibiting high frequencies within the population may be propagated through the population via the use of popular elite sires for breeding [26] or pleiotropic effects linked with enhanced performance in target traits of breeding programs [2]. In this study, the sire with the largest number of offspring (82 animals) carried 4 potential lethal haplotypes, which may have contributed to spreading these haplotypes across generations. Kadri et al. [31] reported the effects of a pleiotropic recessive allele in Nordic Red Dairy cattle that cause embryonic death but is also associated with greater milk yield.

Some of the identified haplotypes are located in adjacent segments with a similar number of carriers, such as haplotypes 41.4, 42.4, 43.4, 44.4, and 46.4 on BTA1; 223.10, 224.9, 225.11 on BTA2; 1736.1 and 1739.3 on BTA13; 1946.13, 1947.12, and 1948.12 on BTA15; 2213.8, 2214.9, and 2214.4 on BTA17; 2291.5 and 2293.2 on BTA18; 2850.17, 2852.5, and 2852.58 on BTA26; and 2922.8 and 2923.7 on BTA27 (Table 8). This result suggests that these hotspots may be influenced by the same lethal variant, reinforcing the potential lethality of these regions. This trend was also observed by Wu et al. [26] and Schmidt et al. [5], where potential lethal haplotypes located in nearby segments were also found.

These findings provide insights into genomic regions that may harbor putative haplotypes containing lethal alleles. Understanding the distribution of these haplotypes within the Nellore breed, as well as identifying their carriers, is of paramount importance to prevent mating between animals carrying the same copy of a lethal allele in a specific locus. This strategy may enhance genetic gain in populations undergoing artificial selection towards specific breeding goals and contribute to the long-term sustainability of breeding programs.

3.4.2. Identified genes and functional enrichment

Considering all genomic regions harboring putative lethal haplotypes, 360 candidate genes were identified (Supplementary File 7, Table S3). The functional enrichment analysis revealed biological processes associated with the development of sexual traits in both males and females, immune system essential functions, energy metabolism, and embryonic development (Table 9). The annotated metabolic pathways included ovarian steroidogenesis, oocyte meiosis, and insulin secretion (Table 9). In this discussion, the main focus was to understand how the identified biological and metabolic processes could be linked to occurrences of embryonic lethality or reproductive inefficiency, which may help in elucidating the effects associated with lethal haplotypes.

The development of primary sexual traits in males and females play a significant role in the cattle reproductive performance, as they can affect fertility in both sexes. The genes *ZFP42* (BTA27), *GATA3* (BTA3), *LHB* (BTA18), *LRP2* (BTA2), *BAX* (BTA18), and *AKR1C3* (BTA13) were identified as regulators of these processes, and some studies have reported their association with reproductive traits in pigs [32], humans [33], zebrafish [34], and cattle [35]. Sires with underdeveloped gonads can exhibit inferior sperm quality, primarily due to lower concentrations of serum follicle-stimulating hormone (FSH), which in turn affects the endocrine regulation of testicular functions and the number of Sertoli cells [36]. On the other hand, heifers with genital developmental issues during the pubertal period may have lower leptin secretion, leading to decreased ovarian follicle development, and lower ovulation rates [37]. This can also be reflected in progesterone concentrations during gestation, increasing the likelihood of embryo loss [38–40]. These processes can impact reproduction in both sexes and consequently influence the expected frequencies of homozygous haplotypes in live animals due to reproductive inefficiency.

The protein-coding gene *AKR1C4* (BTA13) was found to play a role in the hormone metabolic process, converting PGH_2 to $PGF_{2\alpha}$ [41]. In cattle, this conversion was already associated with embryo development and immune function at the fetal-maternal interface during early pregnancy [42, 43]. Uterine production of $PGF_{2\alpha}$ has a negative impact on ongoing embryonic development from the morula to the blastocyst stage, slowing down embryo growth [42, 44].

The *INS* (BTA29) and *IGF2* (BTA29) genes have been identified and linked to glucose homeostasis and energy reserve metabolic process. The function of the *INS* gene is significantly influenced by the hormone metabolism, being crucial for glucose absorption and maintaining glucose balance [45, 46]. In cows, the *INS* gene also controls insulin receptors in embryos from the zygote to the blastocyst stage. Elevated insulin levels during this period can impair the developmental potential of the embryo, resulting in embryo loss [45]. The *IGF2* gene is known to regulate insulin concentrations and facilitate fetal growth by ensuring the supply of oxygen and nutrients to the fetus through placental circulation, contributing to the process of energy homeostasis [47–49]. Studies have documented a mutation in this gene that results in fetal death, attributable to disruptions in the delivery of oxygen and nutrients to the fetus [49, 50].

The regulation of embryonic development was also reported in the enrichment analysis with the genes *RUVBL2* (BTA18), *NLRP5* (BTA18), *PHLDA2* (BTA29) and *ACTR8* (BTA22) being responsible for this process. During embryonic development in cattle, several factors can influence and lead to embryonic loss [51], including incorrect chromosome segregation during cell division in meiosis [52], genetic incompatibility between the maternal immune system and the embryo [53], hormonal disorders [54], and failures during embryo implantation in the maternal uterus [55]. These findings support the hypothesis that the main effect of lethal alleles may be on the occurrence of embryonic loss in affected homozygous embryos [6], possibly due to alterations in the regulation of embryonic development.

The *RUVBL2* gene directly controls transcription through chromatin remodeling in multi-protein complexes [56], and it has been linked to the differentiation of neuroectoderm during early embryogenesis in mouse [57]. Furthermore, the *NLRP5* gene is a protein coding that has been associated with impaired fertility [58] and age at first calving [59] in cattle, regulation of embryo development via mitochondrial functions in mouse [60], and embryo progression in the early stages in ovine species [61]. The expression of the *PHLDA2* gene has been associated with fetal growth restrictions in cattle [62, 63], leading to adverse perinatal outcomes such as embryo death [49] and placental inefficiency [64]. The ACTR8 gene is involved in chromatin remodeling, transcription regulation and DNA recombination [65]. *ACTR8* has been associated with Stayability in Nellore cattle [5] and signatures of selection in Holstein cattle [66].

The metabolic pathways identified in the KEGG enrichment analysis for the gene network were ovarian steroidogenesis (genes *INS*, *LHB* and *AKR1C3*), oocyte meiosis (genes *INS*, *PPP2R1B*, *YWHAH* and *CALM3*), and insulin secretion (genes *INS*, *TRPM4* and *CACNA1D*). These processes are associated with reproductive efficiency in cattle, which supports the hypothesis that the chromosomal regions harboring putative lethal haplotypes may have an impact on the reproductive performance of their carriers.

Ovarian steroidogenesis is a critical process for normal uterine function, as well as the establishment and maintenance of pregnancy [67]. Challenges in this process can also result in the development of nonviable oocytes [67–69]. Conversely, oocyte meiosis in mammals initiates before birth during embryonic development and resumes during puberty in response to luteinizing hormones during estrus [70, 71]. Issues with the oocyte meiosis mechanism can hinder follicle maturation and delay reproduction [71]. Furthermore, insulin secretion plays a pivotal role in body metabolism and can influence cattle reproduction [68, 72]. Adequate insulin concentrations are necessary for normal follicular steroidogenesis and secretion of estradiol [73].

The *INS* gene was found to be involved in all three pathways described. Some studies have indicated that higher insulin and glucose levels are associated with anestrus in Holstein cows [74], while other authors have reported that hyperinsulinemia and increased plasma insulin concentrations impair oocyte quality and subsequent embryo development in cattle [68]. Regarding the other genes associated with metabolic pathways, few studies linking their expression on reproductive traits in cattle were found. For example, the *LHB* gene, involved in the ovarian steroidogenesis mechanism, did not affected the development of ovarian follicular growth or the number of follicles in crossbreed Holstein-Gyr female fetuses, which are crucial factors in ovarian reserve and fertility [75]. On the other hand, the *PPP2R1B* gene was identified in a co-expression network analysis of the metabolome and transcriptome to influence fertility in beef heifers [76].

3.4.3. Reproductive performance of mating involving carriers of lethal haplotypes

In this study, 5093 reproductive records were analyzed, and 45 potential lethal haplotypes were identified, totaling 229,185 possible results for matings between noncarriers or carriers of lethal haplotypes (Supplementary File 9, Table S5). A total of

2564 combinations between the mating class C_{Dam} × C_{Sire} were effectively observed in the reproductive records (Supplementary File 8, Table S3). Among the 1258 genotyped cows that entered reproduction stage, 1209 were identified as carriers of at least one candidate lethal haplotype, while 186 out of 190 genotyped bulls were also found to carry at least one candidate lethal haplotype.

This finding is crucial in demonstrating that despite the low probability of deleterious mutations emerging in populations and being considered rare [2, 26], the lethal haplotypes are prevalent among the animals engaged in reproduction within the evaluated population. This result highlights the importance of implementing targeted mating schemes to prevent combinations between animals harboring identical copies of a candidate lethal haplotype. Furthermore, it is imperative to emphasize that the effective population size is small (107 ± 15 animals), have an estimated founder population of 69 animals, and the breeding program is closed for selecting progenitors exclusively from within each selection line. This scenario difficult the management of the dissemination of these haplotypes within the population, particularly if a bull deemed superior for target traits is a carrier of a candidate haplotype.

Three haplotypes (42.4 – BTA1, 935.15 – BTA6, and 2214.4 – BTA17) had significant effects (P-value < 0.05) on the probabilities of pregnancy success, pregnancy loss, stillbirth, and pre-weaning mortality (Tables 10 and 11). In the mating class C_{Dam} × C_{Sire}, the odds ratio of pregnancy success was 55.14%, 46.18%, and 56.54% lower to NC_{Dam} × NC_{Sire} considering the haplotypes 42.4, 935.15, and 2214.4, respectively (Table 10). The probabilities of observing pregnancy loss in C_{Dam} × C_{Sire} were 50.91% higher when considering the haplotype 42.4, 210.81% higher for haplotype 935.15, and 84.72% higher for haplotype 2214.4 (Table 11). The stillbirth and pre-weaning mortality occurrences were also higher for the C_{Dam} × C_{Sire} mating category (Table 11), with probabilities of 141.02% (stillbirth) and 44.03% (pre-weaning mortality) for haplotype 42.4, 295.03% (stillbirth), and 189.56% (pre-weaning mortality) for haplotype 935.15, and 54.01% (stillbirth) and 101.14% (pre-weaning mortality) for haplotype 2214.4.

These results demonstrate that even in heterozygous carriers of potential candidates for lethal haplotypes, there is a significant decline in their reproductive performance. Similar results were observed in mating categories C_{Dam}× NC_{Sire} and NC_{Dam} × C_{Sire}, where a loss in reproductive performance was evident for certain haplotypes (Tables 10 and 11). This trend has also been reported in other studies. For

instance, Id-Lahoucine et al. [4, 26] found in Angus cattle that the probability of pregnancy decreased by up to 15% in carriers of specific lethal haplotypes. Wu et al. [26] observed two candidate lethal haplotypes in Nordic Holsteins that increased the return-to-estrus rate in previously identified pregnant cows, suggesting early embryonic lethality. Similarly, Id-Lahoucine et al. [77] noted that the occurrence of certain genomic regions with evidence of transmission ratio distortion phenomenon, potentially harboring lethal alleles, increased the likelihood of stillbirth by up to 254%. Likewise, Schmidt et al. [5] identified 15 candidate lethal haplotypes in Nellore cattle that exerted significant effects on post-natal mortality occurrence.

Specifically, considering pregnancy success probabilities, reductions of up to 203.46% in the odds ratio were observed in the mating category $C_{Dam} \times C_{Sire}$ compared to the scenario $NC_{Dam} \times NC_{Sire}$ (Table 10), considering haplotype 998.14 (BTA7). This haplotype was also identified by Schmidt et al. [5], who reported significant negative effects of its occurrence on the probability of heifer rebreeding and stayability. This result supports the finding that this genomic region likely harbors a specific lethal allele in Nellore cattle that influences the reproductive efficiency of its carriers. No annotated genes were overlapped with this haplotype region (36.55 – 37.25 Mb, BTA7).

The haplotype 1739.3 was responsible for the higher increase in the occurrence of pregnancy loss (275.15%) in the mating category $C_{Dam} \times C_{Sire}$ compared to the category $NC_{Dam} \times NC_{Sire}$ (Table 11). The genomic region harboring haplotype 1739.3 (15.76 – 16.49 Mb, BTA13) had not been previously reported as a potential carrier of lethal haplotype candidates in cattle breeds. The genes *GATA3*, *TAF3*, *ATP5F1C*, *KIN*, *ITIH2*, *ITIH5*, and *SFMBT2* were found in this region. It is important to highlight the *GATA3* gene, which was observed in the functional enrichment analysis of the gene network and was associated with biological processes related to development of primary male sexual characteristics, male gonad development, and hormone metabolic processes (Table 9). In cattle, the *GATA3* gene has been linked to regulating lineage specification during early embryonic development, and its deletion disrupts transcription in bovine blastocysts [78].

Mating involving the $C_{Dam} \times C_{Sire}$ category, considering carriers of haplotype 935.15, experienced a 295.03% increase probability in stillbirth occurrence (Table 11). This haplotype is located on BTA6 between 111.03 – 111.68 Mb, encompassing the genes *CD38*, *FGFBP1*, *PROM1*, and *TAPT1*. This genomic region had not been previously reported as a potential host of putative lethal haplotypes. Among the

identified genes, CD38 was the only gene associated with postnatal survival traits in livestock species, with a significant effect reported on the occurrence of stillborn piglets in purebred Yorkshire pigs [79].

Regarding pre-weaning mortality, a 301.40% increase in the probability of its occurrence was observed among matings in the $C_{Dam} \times C_{Sire}$ category, considering carriers of the haplotype 1110.12 (Table 11), compared to the non-carrier category ($NC_{Dam} \times NC_{Sire}$). Among the genes identified in the region of this putative lethal haplotype (21.62 – 22.74 Mb, BTA8), the *CDKN2A* gene was found. This gene has been associated with cell cycle regulation in humans [80] and ribeye area in Nellore cattle [81], providing limited information on the mechanisms of this genomic region that may be influencing the occurrence of pre-weaning mortality.

Taken together, these findings support the argument that lethal alleles in Nellore cattle breeding should be traced within the population to prevent reductions in pregnancy rates, pregnancy loss, stillbirth, and pre-weaning mortality occurrences. Individuals selected for reproduction should be tested for the lethal haplotypes they carry, and mating between individuals carrying the same copy of a lethal haplotype should be avoided. These practices can prevent the spread of lethal alleles to future generations and potentially increase reproductive efficiency, as demonstrated by the results of this study in matings within the $C_{Dam} \times C_{Sire}$ category when compared to the $NC_{Dam} \times NC_{Sire}$ category.

Further studies on this topic involving Nellore cattle are still needed, given the impact of this breed on the socioeconomic landscape and the production of beef for the global market. It is worth noting that among the 45 candidate lethal haplotypes found in this study, only 3 had been reported in previous studies [5]. Although the population evaluated in this study belongs to an experimental herd, there may still be genomic regions containing lethal alleles that have not yet been identified in Nellore cattle. Additional studies utilizing larger databases and whole-genome sequence data would be beneficial to confirm the potential lethality of these haplotypes and reinforcing that their complete absence is not merely a random event. This comprehensive approach would provide more robust evidence regarding the association between these haplotypes, lethal outcomes and reproductive inefficiency, enhancing our understanding of their genetic effects and informing breeding practices for improved livestock management.

3.5. Conclusions

A total of 45 candidate lethal haplotypes were identified in closed experimental Nellore cattle lines. Within the genomic regions of these candidate lethal haplotypes, 360 genes were found, and the enrichment analyses revealed associations with reproductive, hormonal, and metabolic aspects. Carriers of the potential lethal haplotypes exhibited reduced reproductive efficiency compared to noncarriers. This resulted in decreased pregnancy success rates, increased pregnancy loss, stillbirth occurrence, and pre-weaning mortality. These results highlight the importance of monitoring lethal genetic variants in Nellore cattle that exert negative effects on reproduction and production efficiency, avoiding their propagation in the population.

3.6. Supplementary files

All supplementary materials have been deposited in the Harvard Dataverse repository and are available at the permanent link: https://doi.org/10.7910/DVN/T9HCLU

3.7. Declarations

3.7.1. Ethics approval and consent to participate

Animal Care and Use Committee approval was not obtained for this study because all the analyses were performed using pre-existing datasets.

3.7.2. Consent for publication

Not applicable

3.7.3. Availability of data and materials

The haplotypes results obtained in this study have been deposited in the Harvard Dataverse and can be accessed via the following link: https://doi.org/10.7910/DVN/XBAHMZ. The parameters and guidelines used for haplotype identification can be found in the findhap.f90 manual provided by the Agricultural Research Service (USDA). Phenotypic and genomic information can be required for academic use contacting MEZM (email: mezmercadante@gmail.com).

3.7.4. Competing interests

The authors declare that they have no competing interests.

3.7.5. Funding

São Paulo Research Foundation (FAPESP) for financial support (grants 2017/10630-2, 2018/20026-8, and 2017/50339-5), and scholarships to GRDR (grants 2023/11176-4 and 2024/01909-7), LFMM (grant 2022/11852-7), PIS (2021/09942-5), and JPSV (2024/05697-4). Coordination for the Improvement of Higher Education Personnel (CAPES, Finance Code 001) for providing scholarship to ESO. National Council for Science and Technological Development (CNPq) for providing grants to LGA and MEZM.

3.7.6. Authors' contributions

MEZM coordinated the team and supervised all the stages of the study. GRDR, JNSGC, LFMM, and MEZM conceived and designed the study. GRDR, JNSGC, LFMM, PIS, and MEZM conducted the data analyses. LGA and MEZM obtained the resources for this research. GRDR, JNSGC, LFMM, PIS, JPSV, ESO, LGA, LFB, and MEZM contributed to the data acquisition and interpretation of the results. GRDR, JNSGC, LFMM, LFB, and MEZM wrote and edited the manuscript. All authors reviewed and contributed to the editing of the manuscript and approved its final version.

3.7.7. Acknowledgements

The authors thank Dr. Paul M. VanRaden (USDA) for generously providing the codes for obtaining the lethal haplotypes, and the staff of the Institute of Animal Science (Sertãozinho, SP, Brazil) for their efforts in collecting the data used in this study.

References

1. Gebreyesus G, Sahana G, Christian Sørensen A, Lund MS, Su G. Novel approach to incorporate information about recessive lethal genes increases the accuracy of genomic prediction for mortality traits. Hered. 2020;125:155–66.
2. Jenko J, McClure MC, Matthews D, McClure J, Johnsson M, Gorjanc G, et al. 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. Genet Sel Evol. 2019;51:1–14.
3. Wu X, Mesbah-Uddin M, Guldbrandtsen B, Lund MS, Sahana G. Haplotypes responsible for early embryonic lethality detected in Nordic Holsteins. J Dairy Sci. 2019;102:11116–23.
4. Id-Lahoucine S, Casellas J, Lu D, Sargolzaei M, Miller S, Cánovas A. Distortion of Mendelian segregation across the Angus cattle genome uncovering regions affecting reproduction. Sci Rep 2023 13:1. 2023;13:1–10.
5. Schmidt PI, Mota LFM, Fonseca LFS, dos Santos Silva DB, Frezarim GB, Arikawa LM, et al. Identification of candidate lethal haplotypes and genomic association with post-natal mortality and reproductive traits in Nellore cattle. Sci Rep. 2023;13:1–12.
6. VanRaden PM, Olson KM, Null DJ, Hutchison JL. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. J Dairy Sci. 2011;94:6153–61.
7. Hoff JL, Decker JE, Schnabel RD, Taylor JF. Candidate lethal haplotypes and causal mutations in Angus cattle. BMC Genomics. 2017;18:1–11.
8. Benfica LF, Brito LF, do Bem RD, de Oliveira LF, Mulim HA, Braga LG, et al. Detection and characterization of copy number variation in three differentially-selected Nellore cattle populations. Front Genet. 2024;15:1–14.
9. Rodrigues GRD, Rezende VT, Mercadante MEZ, Bonilha SFM, Canesin RC, Raineri C, et al. Animal growth models as a tool to estimate resilience indicators in Bos indicus and Bos taurus heifers: Selection effects and genetics parameters. Livest Sci. 2024;282:1–9.
10. Strandén I. RelaX2 program for pedigree analysis, User's guide for version 1.65. 2014.

11. Gutiérrez JP, Cervantes I, Molina A, Valera M, Goyache F. Individual increase in inbreeding allows estimating effective sizes from pedigrees. Genet Sel Evol. 2008;40:359–78.
12. Sargolzaei M, Chesnais JP, Schenkel FS. A new approach for efficient genotype imputation using information from relatives. BMC Genomics. 2014;15:1–12.
13. Mota LFM, Santos SWB, Júnior GAF, Bresolin T, Mercadante MEZ, Silva JAV, et al. Meta-analysis across Nellore cattle populations identifies common metabolic mechanisms that regulate feed efficiency-related traits. BMC Genomics. 2022;23:1–12.
14. Mota LFM, Arikawa LM, Santos SWB, Fernandes Júnior GA, Alves AAC, Rosa GJM, et al. Benchmarking machine learning and parametric methods for genomic prediction of feed efficiency-related traits in Nellore cattle. Sci Rep. 2024;14:1–14.
15. Misztal I, Tsuruta S, Lourenco D, Masuda Y, Aguilar I, Legarra A, et al. Manual for BLUPF90 family of programs. 1st edition. Athens: University of Georgia; 2014.
16. VanRaden PM, Sun C, O'Connell JR. Fast imputation using medium or low-coverage sequence data. BMC Genet. 2015;16:1–12.
17. Li Y, Sung WK, Liu JJ. Association mapping via regularized regression analysis of single-nucleotide-polymorphism haplotypes in variable-sized sliding windows. Am J Hum Genet. 2007;80:705–15.
18. Guo Y, Li J, Bonham AJ, Wang Y, Deng H. Gains in power for exhaustive analyses of haplotypes using variable-sized sliding window strategy: A comparison of association-mapping strategies. Eur J Hum Genet. 2009;17:785–92.
19. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience. 2020;9:1–9.
20. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation. 2021;2:1–9.
21. R Core Team. R: The R Project for Statistical Computing. 2022. https://www.r-project.org/.
22. McFadden A, Vierra M, Robilliard H, Martin K, Brooks SA, Everts RE, et al. Population Analysis Identifies 15 Multi-Variant Dominant White Haplotypes in Horses. Animals. 2024;14:1–12.

23. Todd ET, Thomson PC, Hamilton NA, Ang RA, Lindgren G, Viklund Å, et al. A genome-wide scan for candidate lethal variants in Thoroughbred horses. Sci Rep. 2020;10:1–10.
24. Ben Braiek M, Fabre S, Hozé C, Astruc JM, Moreno-Romieux C. Identification of homozygous haplotypes carrying putative recessive lethal mutations that compromise fertility traits in French Lacaune dairy sheep. Genet Sel Evol. 2021;53:1–13.
25. Häfliger IM, Spengeler M, Seefried FR, Drögemüller C. Four novel candidate causal variants for deficient homozygous haplotypes in Holstein cattle. Sci Rep. 2022;12:1–13.
26. Wu X, Mesbah-Uddin M, Guldbrandtsen B, Lund MS, Sahana G. Novel haplotypes responsible for prenatal death in Nordic Red and Danish Jersey cattle. J Dairy Sci. 2020;103:4570–8.
27. Sasaki S, Watanabe T, Ibi T, Hasegawa K, Sakamoto Y, Moriwaki S, et al. Identification of deleterious recessive haplotypes and candidate deleterious recessive mutations in Japanese Black cattle. Sci Rep. 2021;11:1–13.
28. Howard JT, Pryce JE, Baes C, Maltecca C. Invited review: Inbreeding in the genomics era: Inbreeding, inbreeding depression, and management of genomic variability. J Dairy Sci. 2017;100:6009–24.
29. Bosse M, Megens HJ, Derks MFL, de Cara ÁMR, Groenen MAM. Deleterious alleles in the context of domestication, inbreeding, and selection. Evol Appl. 2019;12:6–17.
30. Pausch H, Schwarzenbacher H, Burgstaller J, Flisikowski K, Wurmser C, Jansen S, et al. Homozygous haplotype deficiency reveals deleterious mutations compromising reproductive and rearing success in cattle. BMC Genomics. 2015;16:1–13.
31. Kadri NK, Sahana G, Charlier C, Iso-Touru T, Guldbrandtsen B, Karim L, et al. A 660-Kb Deletion with Antagonistic Effects on Fertility and Milk Production Segregates at High Frequency in Nordic Red Cattle: Additional Evidence for the Common Occurrence of Balancing Selection in Livestock. PLoS Genet. 2014;10:1–11.
32. Liu T, Wu Y, Li L, Zhang T, Zhang X, Li Y. High Expression of ZFP42 Improves Early Development of Pig Embryos Produced by Handmade Cloning. Cell Reprogram. 2024;26:57–67.

33. Beiraghdar M, Beiraghdar M, Khosravi S. The methylation status of GATA3 potentially predicts the outcomes of assisted reproductive technologies. Hum Fertil. 2023;26:1279–85.
34. Chu L, Li J, Liu Y, Hu W, Cheng CHK. Targeted gene disruption in zebrafish reveals noncanonical functions of lh signaling in reproduction. Mol Endocrinol. 2014;28:1785–95.
35. Yang MY, Rajamahendran R. Expression of Bcl-2 and Bax proteins in relation to quality of bovine oocytes and embryos produced in vitro. Anim Reprod Sci. 2002;70:159–69.
36. Burns BM, Gazzola C, Holroyd RG, Crisp J, Mcgowan MR. Male reproductive traits and their relationship to reproductive traits in their female progeny: A systematic review. Reprod Domest Anim. 2011;46:534–53.
37. Williams GL, Amstalden M, Garcia MR, Stanko RL, Nizielski SE, Morrison CD, et al. Leptin and its role in the central regulation of reproduction in cattle. Domest Anim Endocrinol. 2002;23:339–49.
38. Carter F, Forde N, Duffy P, Wade M, Fair T, Crowe MA, et al. Effect of increasing progesterone concentration from Day 3 of pregnancy on subsequent embryo survival and development in beef heifers. Reprod Fertil Dev. 2008;20:368–75.
39. Kasimanickam RK, Kasimanickam VR, Oldham J, Whitmore M. Cyclicity, estrus expression and pregnancy rates in beef heifers with different reproductive tract scores following progesterone supplementation. Theriogenology. 2020;145:39–47.
40. Martins KR, Haas CS, Rovani MT, Moreira F, Goetten ALF, Ferst JG, et al. Regulation and function of leptin during ovarian follicular development in cows. Anim Reprod Sci. 2021;227:1–12.
41. Parent M, Madore E, MacLaren LA, Fortier MA. 15-hydroxyprostaglandin dehydrogenase in the bovine endometrium during the oestrous cycle and early pregnancy. Reproduction. 2006;131:573–82.
42. Emond V, Fortier MA, Murphy BD, Lambert2 RD. Prostaglandin E 2 Regulates Both Interleukin-2 and Granulocyte-Macrophage Colony-Stimulating Factor Gene Expression in Bovine Lymphocytes'. Biol Reprod. 1998;58:143–51.
43. Scenna FN, Edwards JL, Rohrbach NR, Hockett ME, Saxton AM, Schrick FN. Detrimental effects of prostaglandin F2α on preimplantation bovine embryos. Prostaglandins Other Lipid Mediat. 2004;73:215–26.

44. Sponchiado M, Gomes NS, Fontes PK, Martins T, Del Collado M, Pastore ADA, et al. Pre-hatching embryo-dependent and -independent programming of endometrial function in cattle. PLoS One. 2017;12:1–23.
45. Laskowski D, Sjunnesson Y, Humblot P, Andersson G, Gustafsson H, Båge R. The functional role of insulin in fertility and embryonic development-What can we learn from the bovine model? Theriogenology. 2016;86:457–64.
46. D'Occhio MJ, Campanile G, Baruselli PS, Porto Neto LR, Hayes BJ, Collins A, et al. Pleomorphic adenoma gene (PLAG1) in reproduction and implication for embryonic survival in cattle: A review . J Anim Sci. 2024;102:1–12.
47. Thatcher' WW, Guzeloglu1 A, Meikle2 A, Kamimura3 S, Bilby' T, Kowalski' AA, et al. Regulation of embryo survival in cattle. Reproduction. 2003;61:253–66.
48. Gobikrushanth M, Purfield DC, Colazo MG, Wang Z, Butler ST, Ambrose DJ. The relationship between serum insulin-like growth factor-1 (IGF-1) concentration and reproductive performance, and genome-wide associations for serum IGF-1 in Holstein cows. J Dairy Sci. 2018;101:9154–67.
49. Sigdel A, Bisinotto RS, Peñagaricano F. Genes and pathways associated with pregnancy loss in dairy cattle. Sci Rep. 2021;11:1–11.
50. Bergman D, Halje M, Nordin M, Engström W. Insulin-like growth factor 2 in development and disease: A mini-review. Gerontology. 2013;59:240–9.
51. Lonergan P, Fair T, Forde N, Rizos D. Embryo development in dairy cattle. Theriogenology. 2016;86:270–7.
52. De Coster T, Masset H, Tšuiko O, Catteeuw M, Zhao Y, Dierckxsens N, et al. Parental genomes segregate into distinct blastomeres during multipolar zygotic divisions leading to mixoploid and chimeric blastocysts. Genome Biol. 2022;23:1–29.
53. Fair T. Embryo maternal immune interactions in cattle. Anim Reprod. 2016;13:346–54.
54. Ribeiro ES, Carvalho MR. Impact and mechanisms of inflammatory diseases on embryonic development and fertility in cattle. Anim Reprod. 2017;14:589–600.
55. Lonergan P, Forde N. Maternal-embryo interaction leading up to the initiation of implantation of pregnancy in cattle. Animal. 2014;8:64–9.
56. Gorynia S, Bandeiras TM, Pinho FG, McVey CE, Vonrhein C, Round A, et al. Structural and functional insights into a dodecameric molecular machine - The RuvBL1/RuvBL2 complex. J Struct Biol. 2011;176:279–91.

57. Hong S, Jo J, Kim HJ, Lee JE, Shin DH, Lee SG, et al. RuvB-Like protein 2 (Ruvbl2) has a role in directing the neuroectodermal differentiation of mouse embryonic stem cells. Stem Cells Dev. 2016;25:1376–85.
58. Pennetier S, Perreau C, Uzbekova S, Thélie A, Delaleu B, Mermillod P, et al. MATER protein expression and intracellular localization throughout folliculogenesis and preimplantation embryo development in the bovine. BMC Dev Biol. 2006;6:1–9.
59. Alves AAC, da Costa RM, Fonseca LFS, Carvalheiro R, Ventura RV, Rosa GJ de M, et al. A Random Forest-Based Genome-Wide Scan Reveals Fertility-Related Candidate Genes and Potential Inter-Chromosomal Epistatic Regions Associated With Age at First Calving in Nellore Cattle. Front Genet. 2022;13:1–18.
60. Fernandes R, Tsuda C, Perumalsamy AL, Naranian T, Chong J, Acton BM, et al. NLRP5 mediates mitochondrial function in mouse oocytes and embryos. Biol Reprod. 2012;86:1–10.
61. Bebbere D, Ariu F, Bogliolo L, Masala L, Murrone O, Fattorini M, et al. Expression of maternally derived KHDC3, NLRP5, OOEP and TLE6 is associated with oocyte developmental competence in the ovine species. BMC Dev Biol. 2014;14:1–13.
62. Guillomot M, Taghouti G, Constant F, Degrelle S, Hue I, Chavatte-Palmer P, et al. Abnormal Expression of the Imprinted Gene Phlda2 in Cloned Bovine Placenta. Placenta. 2010;31:482–90.
63. Wang M, Li D, Zhang M, Yang W, Wu G, Cui Y, et al. Biallelic expression of Tssc4, Nap1l4, Phlda2 and Osbpl5 in adult cattle. J Genet. 2015;94:391–5.
64. Chavatte-Palmer P, Camous S, Jammes H, Le Cleac'H N, Guillomot M, Lee RSF. Review: Placental perturbations induce the developmental abnormalities often observed in bovine somatic cell nuclear transfer. Placenta. 2012;26:99–104.
65. Choe SH, Park SJ, Cho HM, Park HR, Lee JR, Kim YH, et al. A single mutation in the ACTR8 gene associated with lineage-specific expression in primates. BMC Evol Biol. 2020;20:1–12.
66. Naderi S, Moradi MH, Farhadian M, Yin T, Jaeger M, Scheper C, et al. Assessing selection signatures within and between selected lines of dual-purpose black and white and German Holstein cattle. Anim Genet. 2020;51:391–408.

67. Wood JR, Strauss III JF. Multiple Signal Transduction Pathways Regulate Ovarian Steroidogenesis. Rev Endocr Metab Disord. 2002;3:33–46.

68. Sartori R, Guardieiro MM, Surjus RS, Melo LF, Prata AB, Ishiguro M, et al. Metabolic hormones and reproductive function in cattle. Anim Reprod. 2013;10:199–205.

69. Gareis NC, Huber E, Hein GJ, Rodríguez FM, Salvetti NR, Angeli E, et al. Impaired insulin signaling pathways affect ovarian steroidogenesis in cows with COD. Anim Reprod Sci. 2018;192:298–312.

70. Von Stetina JR, Orr-Weaver TL. Developmental control of oocyte maturation and egg activation in metazoan models. Cold Spring Harb Perspect Biol. 2011;3:1–19.

71. He M, Zhang T, Yang Y, Wang C. Mechanisms of Oocyte Maturation and Related Epigenetic Regulation. Front Cell Dev Biol. 2021;9:1–18.

72. Baruselli PS, Batista EOS, Vieira LM, Sales JN de S, Gimenes LU, Ferreira RM. Intrinsic and extrinsic factors that influence ovarian environment and efficiency of reproduction in cattle. Anim Reprod. 2017;14:48–60.

73. Silva JM, Price CA. Insulin and IGF-I are necessary for FSH-induced cytochrome P450 aromatase but not cytochrome P450 side-chain cleavage gene expression in oestrogenic bovine granulosa cells in vitro. J Endocrinol. 2002;174:499–507.

74. Azarbayejani R, Mohammadsadegh M. Glucose, insulin, and cortisol concentrations and glucose tolerance test in Holstein cows with inactive ovaries. Trop Anim Health Prod. 2021;53:3–9.

75. Weller MMDCA, Fortes MRS, Marcondes MI, Rotta PP, Gionbeli TRS, Valadares Filho SC, et al. Effect of maternal nutrition and days of gestation on pituitary gland and gonadal gene expression in cattle. J Dairy Sci. 2016;99:3056–71.

76. Banerjee P, Rodning SP, Diniz WJS, Dyce PW. Co-Expression Network and Integrative Analysis of Metabolome and Transcriptome Uncovers Biological Pathways for Fertility in Beef Heifers. Metabolites. 2022;12:1–18.

77. Id-Lahoucine S, Cánovas A, Legarra A, Casellas J. Transmission ratio distortion regions in the context of genomic evaluation and their effects on reproductive traits in cattle. J Dairy Sci. 2023;106:7786–98.

78. Shi Y, Hu B, Wang Z, Wu X, Luo L, Li S, et al. Functional role of GATA3 and CDX2 in lineage specification during bovine early embryonic development. Reproduction. 2023;165:325–33.

79. Jonas E, Rydhmer L. Effect of candidate genes for maternal ability on piglet survival and growth. Livest Sci. 2018;207:83–90.

80. Nielsen GP, Burns KL, Rosenberg AE, Louis DN. CDKN2A Gene Deletions and Loss of p16 Expression Occur in Osteosarcomas That Lack RB Alterations. 1998.

81. Fernandes Júnior GA, Costa RB, De Camargo GMF, Carvalheiro R, Rosa † G J M, Baldi F, et al. Genome scan for postmortem carcass traits in Nellore cattle. J Anim Sci. 2016;94:4087–95.

CHAPTER 4: UNRAVELING GENOMIC REGIONS WITH TRANSMISSION RATIO DISTORTION HARBORING PUTATIVE LETHAL ALLELES AND THEIR BIOLOGICAL IMPLICATIONS IN NELLORE CATTLE FROM EXPERIMENTAL SELECTION LINES

Gustavo R. D. Rodrigues[*,†,‡,1], Luiz F. Brito[*], Lucio F. M. Mota[†], Joslaine N. S. G. Cyrillo[‡], João B. Silva Neto[*,†], Milena A. F. Campos[*,§], Lenira El Faro[†,‡,#], Lucia G. Albuquerque[†,#], Maria E. Z. Mercadante[†,‡,#,1]

[*]Department of Animal Sciences, Purdue University, West Lafayette, IN, 47907, USA
[†]School of Agriculture and Veterinary Science, São Paulo State University, Jaboticabal, SP, 14884-900, Brazil
[‡]Beef Cattle Research Center, Institute of Animal Science, Sertãozinho, SP, 14174-000, Brazil
[§]School of Veterinary Medicine and Animal Science, Federal University of Bahia, Salvador, BA, 40110-060, Brazil
[#]National Council for Science and Technological Development (CNPq), Brasilia, DF, 71605-001, Brazil
[1]**Corresponding authors**: grd.rodrigues@unesp.br and mezmercadante@gmail.com

Published in Journal of Animal Science (ISSN: 1525-3163)

Received: 25 February 2025
Accepted: 13 June 2025
Published: 17 June 2025
Corrected and typeset: 18 July 2025

Cite this article: Rodrigues, G.R.D., Brito, L.F., Mota, L.F.M. *et al*. Unraveling genomic regions with transmission ratio distortion harboring putative lethal alleles and their biological implications in Nellore cattle from experimental selection lines. *J Anim Sci* **103**, skaf208 (2025). https://doi.org/10.1093/jas/skaf208

Abstract

Transmission Ratio Distortion (TRD) refers to deviations from expected Mendelian inheritance patterns when alleles from heterozygous parents are transmitted at frequencies different from the expected 50%, influenced by biological mechanisms affecting reproduction. Therefore, this study aimed to: i) characterize genomic regions exhibiting TRD patterns in Nellore cattle (*Bos taurus indicus*); ii) detect TRD regions potentially harboring semi-lethal or lethal alleles; and iii) identify candidate genes and perform functional genomic enrichment analyses to reveal quantitative trait loci (QTL), biological processes, and metabolic pathways associated with TRD regions. TRD analyses were performed on a genomic dataset of 3,351 animals with 612,154 autosomal single nucleotide polymorphism (SNP) markers, aligned to the ARS-UCD1.2 *Bos taurus* genome assembly. The software TRDscan v.2.0 was used to evaluate allelic (overall and parent-specific) and genotypic (additive and dominance) parametrizations of TRD effects using a Bayesian framework. Gene annotation, QTL identification, and functional genomic enrichment were conducted based on the locations of the identified TRD regions. A total of 37,783 SNPs and 174,190 haplotypes exhibiting TRD were identified, corresponding to 1,249 genomic regions distributed across all *Bos taurus* autosomes (802 overall TRD, 191 parent-specific TRD, and 256 genotypic TRD). Among these, 73 allelic TRD regions and 59 genotypic TRD regions were identified as potentially harboring semi-lethal or lethal alleles, with the highest number of underrepresented offspring reaching 1,501 individuals. We identified 2,265 candidate genes based on gene annotation, while functional enrichment analyses enabled the identification of 200 significant Gene Ontology (GO) terms, and 2 pathways involved in embryo development, morphogenesis, and growth regulation (P-value < 0.05). Additionally, significant enrichment of QTL associated with production and reproduction traits was observed within these TRD regions (P-value < 0.05). These findings underscore the critical importance of integrating TRD information into genomic selection strategies to enhance productive efficiency while mitigating the spread of deleterious alleles that adversely affect reproduction and embryo survival in Nellore cattle.

Keywords: beef cattle, embryo development, functional annotation, Mendelian inheritance, Quantitative Trait Loci, reproduction.

4.1. Introduction

Transmission Ratio Distortion (TRD) refers to deviations from expected Mendelian inheritance patterns, when specific alleles are disproportionately represented among offspring (Huang et al., 2013; Casellas et al., 2017; Fishman and Mcintosh, 2019). TRD can result from a variety of biological processes, including meiotic drive, gametic selection, zygotic selection, and embryonic or fetal losses (Huang et al., 2013; Fishman and Mcintosh, 2019; Id-Lahoucine et al., 2023a). These mechanisms do not always involve deviations in allele transmission rates but may instead affect the viability of specific genotypes, ultimately altering the proportion of alleles observed among live individuals. In animal breeding, TRD has garnered increasing interest as it reveals non-Mendelian inheritance mechanisms that may influence reproductive success (Abdalla et al., 2020; Vázquez-Gómez et al., 2020; Id-Lahoucine et al., 2023a; Laseca et al., 2023). The investigation of TRD can provide valuable information about genomic regions associated with critical biological outcomes, such as fertility impairments, embryonic lethality, and the segregation of deleterious alleles (Abdalla et al., 2020; Id-Lahoucine et al., 2023a).

Genomic regions exhibiting TRD are particularly relevant due to their potential to harbor putative semi-lethal or lethal alleles. These alleles disrupt essential biological processes or metabolic pathways, often causing developmental disorders and non-Mendelian inheritance patterns (VanRaden et al., 2011; Amorim et al., 2017; Bosse et al., 2019). Genetic variants with harmful effects produce distinct TRD signals by impairing reproduction and newborn survival, frequently leading to embryonic lethality or reduced fitness in heterozygous or homozygous states (Fishman and Mcintosh, 2019; Abdalla et al., 2020). This issue poses significant challenges in livestock breeding, especially in populations under intense selection pressure, where inbreeding elevates the frequency and homozygosity of deleterious alleles (Wright, 1922; VanRaden et al., 2011; Bosse et al., 2019). Consequently, identifying TRD-associated genomic regions is critical for detecting and managing these alleles before their spread in the population.

In Holstein cattle (*Bos taurus taurus*), Id-Lahoucine et al. (2023b) identified 590 TRD regions, linking specific genomic regions to increased risks of non-pregnancy (27%) and stillbirth events (254%). Similarly, in Angus cattle (*Bos taurus taurus*), Id-Lahoucine et al. (2023c) uncovered 851 genomic regions with evidence of TRD, including 71 regions considered to harbor lethal or semi-lethal alleles. These findings

highlight the substantial impact of TRD regions on reproductive performance, emphasizing the need for exploring this phenomenon in other breeds. However, despite the growing interest in TRD, its occurrence and biological implications in Zebu cattle populations (*Bos taurus indicus*) remain poorly understood. Nellore is the predominant Zebu breed raised for beef production in tropical regions due to its greater resilience to environmental challenges (Mota et al., 2020; Silva Neto et al., 2023; Rodrigues et al., 2024a), but to the best of our knowledge, no studies have explored TRD-associated genomic regions in this breed. Investigating TRD in Nellore cattle presents an opportunity to identify genomic regions harboring semi-lethal or lethal alleles and to elucidate their biological functions.

This study aims to identify genomic regions that exhibit TRD patterns in a Nellore cattle population subjected to 44 years of artificial selection in a closed experimental breeding program design. Specifically, we aim to detect TRD regions that may harbor semi-lethal or lethal alleles, considering both allelic and genotypic parametrizations. Furthermore, candidate genes within these regions will be identified, and functional enrichment analyses will be performed to uncover biological processes and metabolic pathways underlying TRD regions. These findings are expected to contribute to the understanding of TRD in Nellore cattle, laying the foundation for future studies in Zebu breeds and informing strategies to enhance reproductive efficiency and sustainability in livestock production.

4.2. Material and Methods

4.2.1. Ethical statement

Data was obtained from an existing database, and therefore, approval from the Ethics Committee was not required.

4.2.2. Population

All datasets used in this study come from a Nellore cattle population developed and maintained within an experimental breeding program at the Beef Cattle Research Center (Institute of Animal Science – IZ, Sertãozinho, SP, Brazil). This population includes three distinct selection lines subjected to 44 years of genetic selection (1980-

2024), with varying criteria for yearling weight (YW) applied over the decades. The program began in 1980 with the objective of selecting animals based on YW, measured at 378 days of age for young bulls and 550 days of age for heifers.

Three selection lines were established: Line 1—Nellore Control (NeC), in which animals were selected based on YW close to the average of their contemporary group (CG) within the same birth year and herd; Line 2—Nellore Selection (NeS), where animals were selected for higher YW within their birth year and herd; and Line 3—Nellore Traditional (NeT), which involved selecting for higher YW and, since 2008, also selected considering lower estimated breeding values (EBVs) for residual feed intake (RFI) within the same birth year and herd. These selection lines are closed, meaning only bulls and dams born within the respective lines are used for breeding, with exception of NeT, which occasionally received animals from NeS. Inbreeding levels are carefully managed through planned matings based on co-ancestry coefficients (Bem et al., 2024).

The evaluated population holds significant relevance for the Nellore cattle population in Brazil, as the founding animals of the three selection lines were representative sires from the breed's national genetic foundation (Cyrillo et al., 2001; Mercadante et al., 2003). Furthermore, genetic material (semen and embryos) from elite animals (NeS and NeT herds) with superior EBVs for growth, feed efficiency, carcass, and reproductive traits have been commercialized and used for breeding across various Brazilian states and in other Latin American countries. Despite being closed herds under selection for 44 years, the evaluated selection lines remain highly relevant to the genetic improvement of the Nellore breed. Further details regarding the experimental breeding program and the selection lines can be found in Benfica et al. (2024) and Rodrigues et al. (2024a).

4.2.3. Genomic and pedigree data

The pedigree file used for the TRD analyses comprised 13,088 individuals, including 486 sires and 3,007 dams. The genomic dataset consisted of 3,351 animals, of which 1,766 were males and 1,581 were females, genotyped using three single nucleotide polymorphism (SNP) chip assays. The detailed information about the three SNP chips used, as well as the distribution of animals per chip, can be found in Rodrigues et al. (2024b) which utilized the same genotype database. Among these,

2,571 animals genotyped with medium-density SNP panels (50K and 75K SNP panels) were imputed to the Illumina BovineHD panel (770K) using FImpute v3 software (Sargolzaei et al., 2014), with a reference population of 6,862 animals genotyped on the HD SNP chip. Before imputation, genotype quality control was conducted, which involved the exclusion of markers that were i) monomorphic, ii) had duplicated coordinates, iii) were located on non-autosomal chromosomes, and iv) had a GenCall score below 0.90. Genome coordinates for the imputation analyses were based on the ARS-UCD1.2 *Bos taurus* genome assembly (Rosen et al., 2020). The expected accuracy of genotype imputation exceeded 0.97 (Mota et al., 2024). After genotype imputation, the genomic dataset included 612,154 autosomal SNP markers and 3,351 animals that were considered for the TRD analyses. Descriptive statistics related to the pedigree and genomic datasets are presented in Table 12.

Table 12 Descriptive statistics of the pedigree and genomic datasets used for transmission ratio distortion (TRD) analyses in experimental lines of Nellore cattle.

Item	Value
SNPs[1] evaluated	612,154
Individuals in the pedigree	13,088
Individuals genotyped	3,351
Sires genotyped	264
Dams genotyped	1,295
Genotyped Individual–Sire pairs	2,476
Genotyped Individual–Dam pairs	2,212
Genotyped Individual–Sire –Dam trios	1,891

[1]Single nucleotide polymorphism.

4.2.4. Identification of TRD regions

Specifically, the TRD analyses rely on the number of genotyped parent-offspring trios (individual – sire – dam), as the availability of trios can enhance the statistical power for estimating TRD parameters (Id-Lahoucine et al., 2019). Each trio enables the evaluation of whether the transmission of alleles from heterozygous parents deviates from the expected Mendelian ratio, a key aspect in identifying regions of the genome that exhibit TRD patterns. The number of informative offspring, defined as the

number of progenies derived from one or both heterozygous parents, are the basis for detecting deviations, as they allow for the assessment of whether one allele is transmitted more or less frequently than expected. A larger number of trios increases the reliability and accuracy of TRD estimates by minimizing random TRD and enhancing the ability to detect distortions (Id-Lahoucine et al., 2019). In the present study, 1,891 trios were available for analyses (Table 12). Potential TRD regions were identified using two specific approaches: i) allelic parametrization of TRD (Casellas et al., 2014) and ii) genotypic parametrization of TRD (Casellas et al., 2012; Casellas et al., 2020), using both SNP-by-SNP and haplotype-based methods (Id-Lahoucine et al., 2019).

The allelic parameterization of TRD describes the probability (P) of allele transmission from heterozygous parents (A/B) to their offspring. This model incorporates a single overall TRD effect (α) in a parent-unspecific framework or distinguishes between sire-specific (α_s) and dam-specific (α_d) TRD effects in a parent-specific model (Casellas et al., 2014; Casellas et al., 2020). The probabilities for both the parent-unspecific and parent-specific models can be expressed as follows:

$$P(A) = 1 - P(B) = 0.5 + \alpha \quad \text{and} \quad P(B) = 1 - P(A) = 0.5 + \alpha$$

$$P_i(A) = 1 - P_i(B) = 0.5 + \alpha_i \quad \text{and} \quad P_i(B) = 1 - P_i(A) = 0.5 + \alpha_i \text{, with } i = [s, d]$$

where α, α_s and α_d are TRD parameters assuming flat priors within a parametric space ranging from -0.5 to 0.5 (Id-Lahoucine et al., 2019). In a Bayesian inference framework, the conditional posterior probabilities of the TRD parameters can be defined as:

$$p(\alpha|y) \propto (p|\alpha)p(\alpha) \quad \text{and} \quad p(\alpha_s, \alpha_d|y) \propto p(y|\alpha_s, \alpha_d)p(\alpha_s)p(\alpha_d)$$

where y is the vector of genotypes of the offspring generation.

The genotypic parameterization of TRD incorporates both additive (α_g) and dominance (δ_g) parameters, without considering the origin of each allele (Casellas et al., 2012; Casellas et al., 2020). The assumed probability of offspring (P_{off}) resulting from a heterozygous-by-heterozygous mating is as follows:

$$P_{off}(AA) = \frac{(1 + \alpha_g - \delta_g)}{4},\ P_{off}(AB) = \frac{(1 + \delta_g)}{2} \text{ and } P_{off}(BB) = \frac{(1 + \alpha_g - \delta_g)}{4}$$

where α_g and δ_g are additive- and dominance-TRD parameters, respectively. The frequencies in offspring from heterozygous-by-homozygous matings were also obtained to guarantee that P_{off}(AA) + P_{off}(AB) + P_{off}(BB) = 1. Under Bayesian

implementation, the conditional posterior probabilities of the TRD parameters considering the genotypic parametrization were defined as:

$$P(\alpha_g, \delta_g|y) \propto P(y|\alpha_g, \delta_g)P(\alpha_g)p(p\alpha_g, \delta_g)$$

where y is the vector of genotypes of the offspring generation. Priors were assumed for α_g and $\delta_g$considering a parametric space of -1 to 1 (Id-Lahoucine et al., 2019).

The allelic parameterization of TRD ranges from -0.5 to 0.5 as it models deviations from Mendelian expectations in allele transmission from heterozygous parents, reflecting biases in inheritance. In contrast, the genotypic parameterization, ranging from -1 to 1, does not account for parental origin but instead models genotype frequencies directly, incorporating additive and dominance effects to capture both transmission and segregation distortions.

The TRD was evaluated on a SNP-by-SNP basis and across different haplotype sizes (5, 10, 15, 20, and 25 SNPs) considering the sliding windows method (Li et al., 2007; Guo et al., 2009) within the 612,514 autosomal SNPs available for this study. For the haplotype sliding windows analyses, the biallelic haplotypes and heterozygous pairwise combinations procedures were considered (Id-Lahoucine et al., 2019). All analyses were performed within a Bayesian framework using the TRDscan v.2.0 software (Id-Lahoucine et al., 2019), employing a single Monte Carlo Markov chain with 110,000 iterations, where the first 10,000 iterations were discarded as burn-in, and an interval of 10 iterations were sampled to estimate the TRD parameters. The overall (parent-unspecific), parent-specific, and genotypic parameterizations of TRD were compared using the Deviance Information Criterion (DIC) to assess the adjustment of these models and inheritance patterns of SNPs and haplotypes.

TRD events were identified based on quality control criteria outlined by Id-Lahoucine et al. (2019; 2023a; 2023b; 2023c): 1) For each SNP or haplotype window, a minimum number of 20 heterozygous sires, 50 heterozygous dams, and 1,000 informative offspring were required to ensure greater statistical power in the inference regarding the TRD phenomena; 2) TRD regions that presented an approximate empirical null distribution of TRD > 0.001% margin error were discarded to minimize the likelihood of false positives due to random variation (gamete sampling). Details regarding the empirical null distribution parameters can be found in Id-Lahoucine et al. (2019); 3) TRD regions with a coefficient of variation greater than 20% for the TRD effects were excluded to ensure consistency, as regions with a narrower confidence

interval are more reliable; and, 4) regions with fewer than four heterozygous sires exhibiting highly skewed transmission, which could be indicative of genotyping errors, were discarded.

Following these steps, the statistical significance of the TRD effect was evaluated using the Bayes factor (BF; Kass and Raftery, 1995). To be considered as statistically significant, SNPs were required to have a Log_{10} (BF) > 100, and haplotype windows (5, 10, 15, 20, and 25 SNPs) needed to have a Log_{10} BF > 50 for the TRD effect. These thresholds were defined to minimize the risk of false positives and to prioritize the most biologically meaningful TRD signals supported by both the data and statistical evidence. Specifically, the threshold applied corresponded approximately to the top 0.01% of all TRD signals detected, serving as a highly stringent filter to ensure robustness given the limitations of the dataset. A similar strategy has been previously employed in epistatic TRD analyses to focus on the most significant results (Id-Lahoucine et al., 2023d). The threshold values for the BF in SNP-by-SNP and haplotype window analyses differed due to the greater number of parameters and combinations that need to be estimated in the haplotypes analyses, which may result in a lower BF for this approach (Id-Lahoucine et al., 2019). Only TRD regions meeting all these criteria were considered for further analyses.

Since many SNPs or haplotypes could be observed with evidence of TRD events and located close to each other due to high linkage disequilibrium (LD), genomic regions were defined. The LD between the SNPs was calculated using the PLINK software (Purcell et al., 2007), and markers located within 100 Kb (upstream and downstream) of each other on the same chromosome, with an r^2 value greater than 0.80, were grouped and classified as part of the same genomic region. After defining the genomic regions, TRD regions were identified, and the genetic variant (SNP or haplotype window) with the highest BF within each genomic region exhibiting evidence of TRD was selected to represent the TRD event within that region.

4.2.5. Detection of TRD regions harboring potential lethal alleles

After identifying the genomic regions associated with allelic and genotypic TRD patterns, candidate regions potentially harboring lethal or semi-lethal alleles were identified based on the number of underrepresented offspring. In the context of TRD

analyses, underrepresented offspring are those whose expected inheritance pattern is not observed, resulting in a lower frequency of a given allele or genotype than anticipated under Mendelian expectations (Abdalla et al., 2020). These offspring are indicative of a selective disadvantage for the under-transmitted allele, possibly due to lethal or semi-lethal effects that impact the viability of the offspring (Id-Lahoucine et al., 2023c).

For the allelic TRD model, the number of underrepresented offspring ($N_{\text{alellic underrepresented}}$) was calculated as:

$$N_{\text{alellic underrepresented}} = N_{\text{informative}} \times 2 \times |\alpha|$$

where $N_{\text{informative}}$ is the number of informative offspring and α is the estimated allelic TRD effect (overall, sire-specific, or dam-specific TRD effect).

In the genotypic TRD model, underrepresented genotypes were identified by comparing observed genotype counts with those expected under Mendelian inheritance across specific mating types. For instance, the expected genotype proportions under the null hypothesis (no TRD) are: AA × AB = 50% AA_1, 50% AB_1; AB × BB = 50% AB_2, 50% BB_2; and AB × AB = 25% AA_3, 50% AB_3, 25% BB_3

where: AA_1 and AB_1 are the offspring genotypes from AA × AB matings; AB_2 and BB_2 are the offspring genotypes from AB × BB matings; and AA_3, AB_3, and BB_3 are the offspring genotypes from AB × AB matings.

Accordingly, the number of underrepresented offspring ($N_{\text{genotypic underrepresented}}$) in the genotypic model was calculated based on deviations from these expectations using the following expression:

$$N_{\text{genotypic underrepresented}} = (AA_1 - AB_1) + (AB_2 - BB_2) + (2 \times AA_3 - AB_3) + (AA_3 - BB_3)$$

To ensure comprehensive detection of TRD effects, reciprocal mating types (e.g., AA sire × AB dam and AB sire × AA dam) were both considered in the analyses. Genomic regions were prioritized as putative carriers of lethal or semi-lethal alleles based on model-specific thresholds established to capture biologically relevant deviations from Mendelian inheritance. In the allelic TRD model, regions were selected when the estimated TRD effect ($|\alpha|$, $|\alpha_s|$, or $|\alpha_d|$) was ≥ 0.20, and the number of underrepresented offspring exceeded 1,000. For the genotypic TRD model, candidate regions were identified when the additive effect ($|\alpha_g|$) exceeded 0.50, the dominance TRD effect ($|\delta_g|$) exceeded 0.10, and the total number of underrepresented offspring also surpassed 1,000. These thresholds were adopted to ensure the selection of

regions with robust evidence of transmission distortion and potential impacts on embryonic or postnatal viability.

4.2.6. Gene annotation, QTL identification, and functional enrichment analyses

The genomic regions identified as being associated with allelic or genotypic TRD patterns were used for gene annotation and subsequent functional enrichment analyses. Gene annotation and identification of Quantitative Trait Loci (QTL) overlapping with TRD regions were performed using the GALLO R package (Fonseca et al., 2020). For gene annotation, data from the Ensembl database for *Bos taurus* and the ARS-UCD1.2 genome assembly (Rosen et al., 2020) were utilized. The annotation of QTL was performed using data from the Cattle QTL Database (Hu et al., 2022), with a focus on QTL associated with production, fertility, and reproduction traits. This approach contributes to identifying specific TRD regions overlapping with QTLs associated with production efficiency traits under direct selection in the evaluated population (YW and RFI). Such findings could provide evidence for the presence of TRD in genomic regions influencing the phenotypic expression of target traits. Furthermore, QTLs associated with reproductive traits were emphasized due to the critical role of reproduction in shaping non-Mendelian patterns of inheritance and its well-documented potential to drive TRD events (Huang et al., 2013; Fishman and Mcintosh, 2019).

Positional candidate genes were further investigated for their functional profiles through genomic enrichment analyses of Gene Ontology (GO) terms related to biological processes (BP), metabolic functions (MF), and cellular components (CC), as well as Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Functional enrichment analyses were conducted using the gprofiler2 R package (Kolberg et al., 2023). The Bonferroni correction for multiple testing was applied to calculate P-values for gene and QTL enrichment analyses, with a significance threshold set at 0.05.

4.3. Results

4.3.1. Identification of TRD regions

Table 13 summarizes the number of SNPs, haplotypes, and genomic regions associated with TRD based on the established criteria for identifying potential TRD events. A total of 37,783 SNPs exhibited evidence of TRD with a Log_{10} BF > 100, and

174,190 haplotypes had a Log_{10} BF > 50 with different window sizes (comprising 5, 10, 15, 20, or 25 SNPs). The highest Log_{10} BF observed for the overall, parent-specific, and genotypic TRD were 252.43, 160.71, and 145.66, respectively. Manhattan plots depicting the BF values for the overall TRD effect considering SNP-by-SNP and the different haplotype windows size are presented in Figure 5.

Table 13 Single nucleotide polymorphisms (SNPs), haplotypes, and genomic regions associated with transmission ratio distortion (TRD) in Nellore cattle

Parameter	Number
SNPs exhibiting TRD	37,783
Haplotypes exhibiting TRD	174,190
Genomic regions exhibiting TRD	1,249
Overall-TRD	802
Parent-specific TRD	191
Genotypic-TRD	256

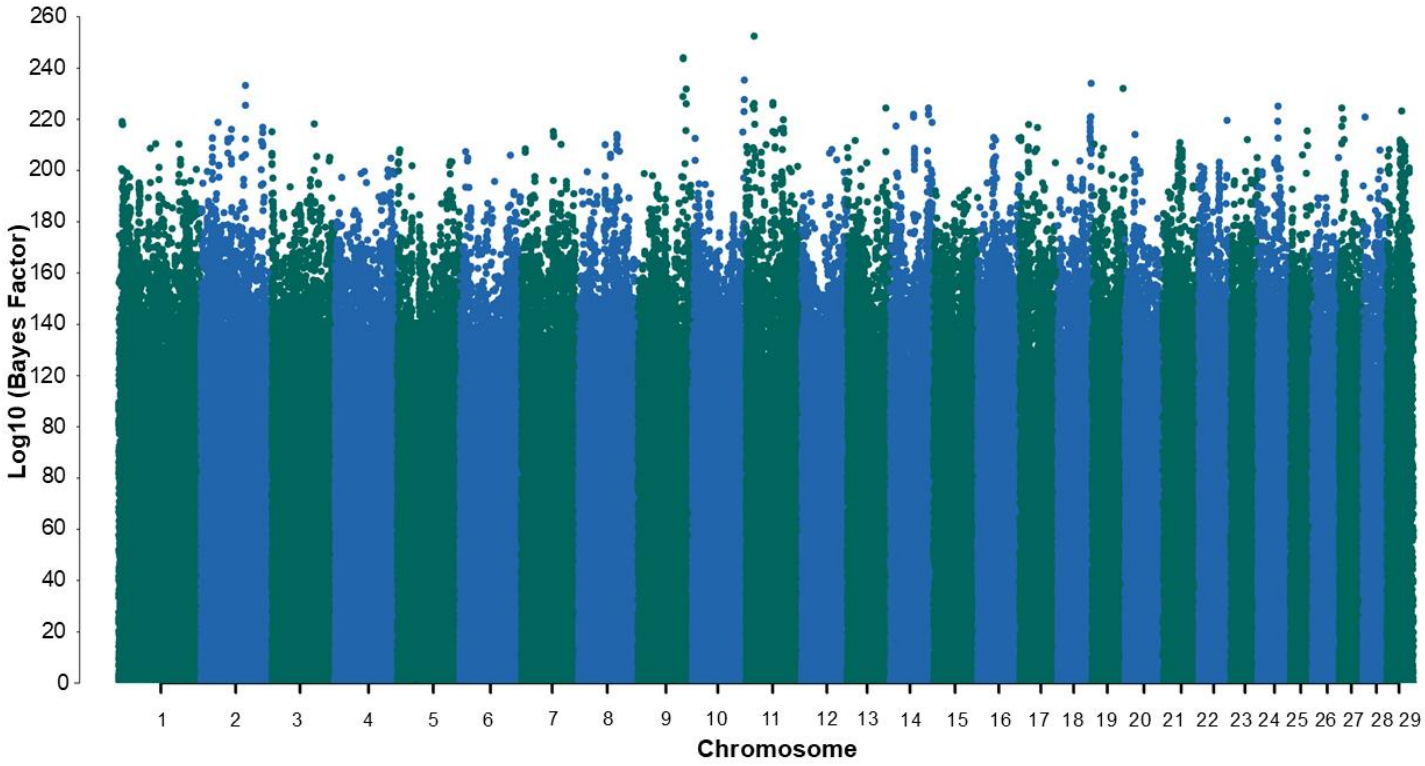

Figure 5. Manhattan plot of the Log_{10} Bayes Factor for the overall transmission ratio distortion (TRD) effect, considering SNP-by-SNP and moving haplotype windows of 5, 10, 15, 20, and 25 SNPs in Nellore cattle.

Considering the SNPs (n = 35,500) and haplotypes (n = 160,231) associated with the overall TRD event, the mean α (± standard deviation) was 0.058 ± 0.001, with the most extreme value observed on BTA10 (100.20Mb) with 0.431 ± 0.007. For SNPs (n = 1,688) and haplotypes (n = 9,784) associated with parent-specific TRD effects, it was observed that in certain genomic regions, sires and dams exhibited opposite directions of TRD effects. This suggests that certain alleles were transmitted preferentially by sires but less so by dams, while others were transmitted preferentially by dams but to a lesser extent by sires. Nevertheless, the largest α_s observed was 0.445 ± 0.010 (BTA9, 95.15 Mb), with an average of 0.086 ± 0.001. For α_d, the largest observed effect was 0.443 ± 0.010 (BTA18, 64.66 Mb), with an average of 0.099 ± 0.012. Regarding the genotypic model of TRD, it was observed that additive effects exhibited a higher Log_{10} BF compared to dominance effects, indicating stronger and more decisive evidence for additive effects over dominance effects. For the SNPs (n = 595) and haplotypes (n = 4,175) where the genotypic parameterization of TRD effects provided a more accurate description of the transmission pattern in the respective regions, the mean α_g was -0.129 ± 0.001, with a minimum value of -0.994 ± 0.001 (BTA24, 43.42 Mb) and a maximum value of 0.992 ± 0.001 (BTA24, 43.41), while the δ_g ranged from -0.196 ± 0.024 (BTA4, 101.76 Mb) to 0.022 ± 0.023 (BTA19, 16.41 Mb) with an average value of -0.110 ± 0.023.

It is important to note that many of the SNPs and haplotype windows showing evidence of TRD were mapped to the same genomic regions. For instance, some regions exhibited consistent TRD patterns where a single SNP, along with all haplotype windows containing that SNP were associated with the same TRD effect. Furthermore, these SNPs and haplotypes were grouped into genomic regions based on LD between adjacent SNPs in the same chromosome. As a result, 1,249 genomic regions were associated with TRD, categorized into three types: overall TRD (802 regions), parent-specific TRD (191 regions), and genotypic TRD (256 regions) (Table 13). None of the TRD regions identified in this study overlapped with those previously reported in *Bos taurus taurus* populations. The genomic regions exhibiting TRD signals covered approximately 9.22% of the bovine genome, corresponding to a total of ~249.8 Mb. Among the parent-specific TRD regions, seven genomic regions exhibited evidence of TRD effects originating from the dam, while 183 regions showed TRD effects associated with the sire. Notably, one region displayed distorted allele transmission from both the sire and dam simultaneously (BTA2, 32.45 – 32.65 Mb). The distribution

of the TRD regions across the *Bos taurus* autosomes (BTAs) can be visualized in Figure 6. In general, all BTAs contained at least one genomic region with a TRD effect, with nearly all chromosomes exhibiting all three types of patterns. BTA25 was the only exception, showing evidence exclusively for overall and genotypic TRD events, without any parent-specific TRD. The comprehensive list detailing the genomic regions associated with TRD and their specific parameters is available in Supplementary File 10, Table S6.

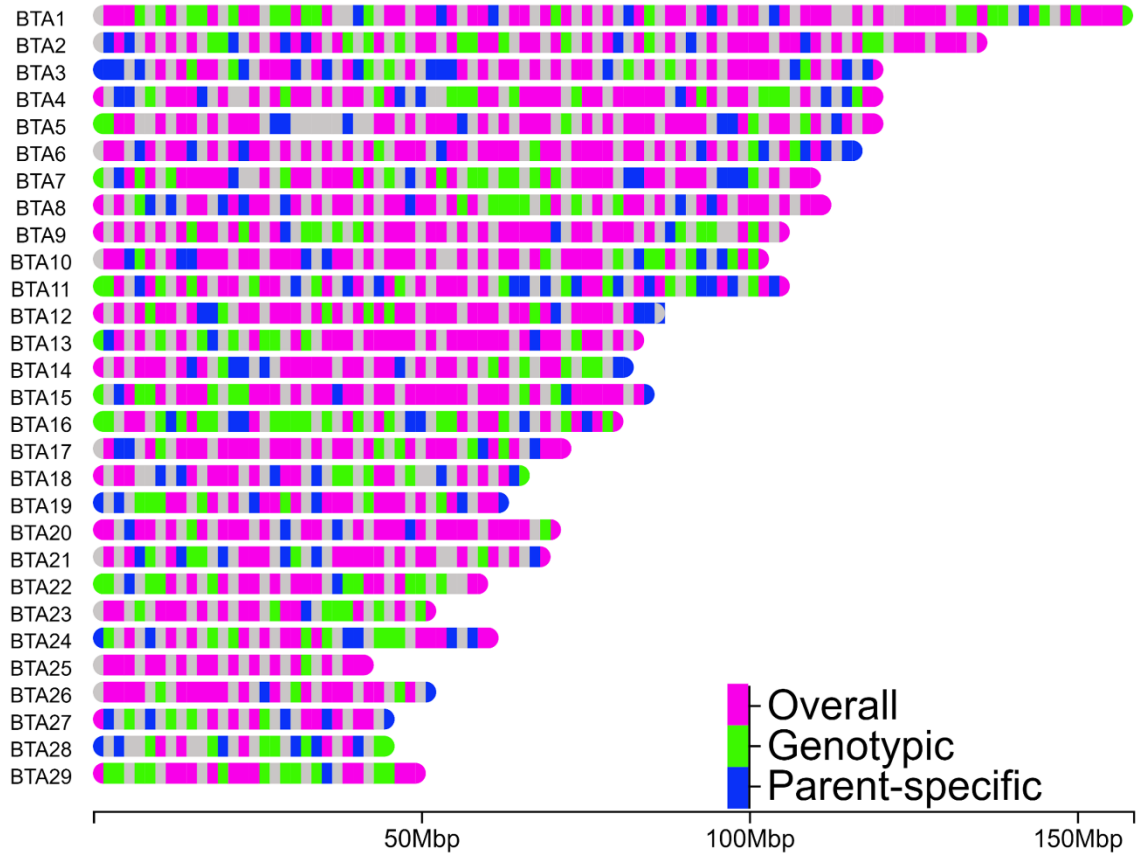

Figure 6. Distribution of transmission ratio distortion regions (overall, parent-specific, and genotypic) by *Bos taurus* autosome (BTA).

4.3.2. Genomic regions with TRD harboring semi-lethal or lethal alleles

The number of underrepresented offspring in genomic regions exhibiting evidence of TRD ranged from 323 to 1,501, with an average of 881 ± 145 individuals. After establishing a threshold of 1,000 underrepresented offspring and a minimum TRD effect of |0.20| to identify potential allelic TRD genomic regions harboring semi-lethal or lethal alleles, 73 regions were identified. Table 14 presents the 30 genomic regions with the highest number of underrepresented offspring, while the complete list containing the 73 candidates is provided in Supplementary File 11, Table S7. Among the 73 regions with allelic TRD patterns, 45 were associated with overall TRD, and 28 with parent-specific TRD (27 sire TRD and 1 dam TRD). The α in these regions ranged from -0.402 ± 0.008 (BTA17, 16.80 – 17.00 Mb) to 0.390 ± 0.008 (BTA14, 10.79 – 10.99 Mb). The α_s ranged from -0.428 ± 0.009 (BTA14, 79.58 – 79.78 Mb) to 0.427 ± 0.009 (BTA24, 38.84 – 39.04 Mb), and the α_d observed was -0.430 ± 0.012 (BTA29, 35.28 – 35.48 Mb). It is noteworthy that, BTA15, BTA16, BTA20, and BTA25 did not exhibit any potential allelic TRD.

Table 14 Genomic regions potentially associated with semi-lethal or lethal alleles based on allelic parametrization of transmission ratio distortion (TRD) in Nellore cattle

Chr[1]	Region[2]	Freq[3]	SireH[4]	DamH[5]	Informative Offspring[6]	TRD type[7]	TRD effects[8]	Log_{10} (BF)[9]	Underrepresented offspring[10]
BTA29	35.28 - 35.48	0.046	45	232	1,572	Parent-Specific (Dam)	α_d = -0.43	103.51	1,374
BTA9	92.70 - 92.90	0.080	82	302	1,720	Parent-Specific (Sire)	α_s = 0.36	144.61	1,249
BTA10	20.55 - 20.75	0.878	100	448	2,096	Overall	α = 0.27	175.89	1,146
BTA11	16.55 - 16.75	0.094	77	338	1,690	Overall	α = -0.34	218.05	1,134
BTA28	3.90 - 4.10	0.929	87	264	1,609	Overall	α = 0.35	220.82	1,132
BTA11	38.48 - 38.68	0.062	48	246	1,408	Parent-Specific (Sire)	α_s = -0.40	138.65	1,128
BTA3	35.19 - 35.39	0.934	73	250	1,448	Parent-Specific (Sire)	α_s = 0.39	141.91	1,119
BTA6	116.48 - 116.68	0.082	61	260	1,503	Parent-Specific (Sire)	α_s = -0.37	132.51	1,118
BTA19	59.76 - 59.96	0.942	64	247	1,493	Overall	α = 0.37	231.96	1,114
BTA17	59.49 - 59.69	0.924	64	291	1,492	Parent-Specific (Sire)	α_s = 0.37	114.66	1,104
BTA18	61.88 - 62.08	0.888	87	403	1,874	Overall	α = 0.29	178.58	1,100
BTA2	21.94 - 22.14	0.068	71	271	1,541	Overall	α = -0.36	212.50	1,097
BTA23	32.46 - 32.66	0.055	62	211	1,442	Parent-Specific (Sire)	α_s = -0.38	149.23	1,092
BTA6	11.49 - 11.69	0.078	73	286	1,588	Overall	α = -0.34	207.29	1,090
BTA1	61.38 - 61.58	0.068	69	250	1,446	Parent-Specific (Sire)	α_s = -0.38	114.48	1,081
BTA1	5.88 - 6.08	0.075	79	294	1,562	Overall	α = -0.34	200.64	1,076
BTA2	44.09 - 44.29	0.887	83	382	1,752	Overall	α = 0.31	183.51	1,074
BTA9	87.33 - 87.53	0.041	51	202	1,261	Parent-Specific (Sire)	α_s = -0.42	157.28	1,069
BTA5	4.57 - 4.77	0.070	76	303	1,524	Overall	α = -0.35	208.04	1,069
BTA3	50.90 - 51.10	0.079	72	314	1,404	Parent-Specific (Sire)	α_s = -0.38	116.54	1,068
BTA8	28.93 - 29.13	0.928	63	257	1,382	Parent-Specific (Sire)	α_s = 0.39	126.89	1,067
BTA11	87.07 - 87.27	0.075	70	302	1,561	Overall	α = -0.34	199.51	1,065

BTA2	34.37 - 34.57	0.077	83	319	1,562	Overall	$\alpha = -0.34$	202.03	1,062
BTA17	68.60 - 68.80	0.922	71	261	1,518	Overall	$\alpha = 0.35$	203.03	1,060
BTA23	28.71 - 28.91	0.075	83	273	1,548	Overall	$\alpha = -0.34$	199.61	1,058
BTA14	44.53 - 44.73	0.943	75	241	1,405	Overall	$\alpha = 0.38$	220.95	1,058
BTA2	42.38 - 42.58	0.065	67	260	1,373	Parent-Specific (Sire)	$\alpha_s = -0.38$	130,21	1,055
BTA18	4.44 - 4.64	0.085	84	293	1,624	Overall	$\alpha = -0.32$	185.46	1,054
BTA1	18.71 - 18.91	0.101	80	347	1,683	Overall	$\alpha = -0.31$	181.05	1,053
BTA14	79.58 – 79.78	0.038	58	166	1,229	Parent-Specific (Sire)	$\alpha_s = -0.43$	157.89	1,053

[1]Chr: Chromosome; BTA: Bos taurus autosome. [2]The region reported is in megabase unit (Mb) and corresponds to the location considering the ARS-UCD1.2 *Bos taurus* genome assembly (Rosen et al., 2020). [3]Frequency of the allele in the population. [4]Number of heterozygous sires. [5]Number of heterozygous dams. [6]Informative offspring were defined as the number of progenies derived from one or both heterozygous parents. [7]Indicate the type of TRD that better describes the inheritance pattern observed. [8]α, α_s, and α_d represent the overall, sire, and dam allelic TRD effect, respectively. [9]The Log_{10} (BF) indicates the Bayes Factor for the TRD effect. [10]The number of underrepresented offspring refers to the total expected offspring that were not observed for a given allele. This value is equivalent to the number of informative offspring multiplied by twice the magnitude of the TRD effect (Id Lahoucine et al., 2020; Id Lahoucine et al., 2023a). Full list containing the 73 genomic regions harboring potential semi-lethal or lethal alleles showing TRD allelic patterns can be found in Supplementary File 11, Table S11.

Considering the genotypic parametrization of TRD models, 59 genomic regions containing potential candidates for semi-lethal or lethal alleles were identified. Table 15 presents the 30 genomic regions with the highest number of underrepresented offspring, while the complete list containing the 59 candidates is provided in Supplementary File 12, Table S8. Within these regions, the higher estimated $|\alpha_g|$ was 0.993 ± 0.006 on BTA24 (43.31 – 43.51 Mb), with strong statistical support for additive TRD events ($\log_{10}$ BF > 100). The highest $|\delta_g|$ was 0.196 ± 0.024 on BTA4 (101.66 – 101.86 Mb), with the strongest evidence for a dominance TRD signal reflected by a $\log_{10}$ BF of 24.64. In some genotypic TRD regions, no AA or BB individuals were observed among the offspring of heterozygous-by-heterozygous (AB × AB) matings. Specifically, the TRD region in which no AA individuals were detected was located on BTA1 (9.54–9.74 Mb). Likewise, regions with complete absence of BB individuals from AB × AB matings were identified on four chromosomes: BTA1 (40.84–41.04 Mb), BTA13 (1.24–1.44 Mb), BTA24 (43.31–43.51 Mb), and BTA27 (26.17–26.37 Mb).

Table 15 Genomic windows potentially associated with semi-lethal or lethal alleles based on genotypic parametrization of transmission ratio distortion (TRD) in Nellore cattle

Chr[1]	Region[2]	Freq[3]	SireH[4]	DamH[5]	Informative Offspring[6]	TRD effects[7]				Number of underrepresented offspring[8]
						α_g	$\text{Log}_{10}(BF_{\alpha_g})$	δ_g	$\text{Log}_{10}(BF_{\delta_g})$	
24	43.31 – 43.51	0.8435	116	459	2,145	0.993	145.66	-0.125	12.26	1,501
4	101.66 –101.86	0.8625	113	396	2,054	0.991	105.19	-0.196	24.64	1,368
19	41.63 -41.83	0.8349	120	449	2,079	0.990	131.48	-0.111	12.65	1,361
1	40.84 -41.04	0.8470	104	451	2,074	0.991	120.50	-0.129	11.42	1,344
13	1.24 -1.44	0.8544	111	428	2,023	0.989	110.57	-0.162	18.76	1,340
7	61.87 -62.07	0.8344	104	452	2,005	0.989	116.84	-0.158	23.82	1,327
18	64.52 -64.72	0.8463	97	462	2,041	0.986	108.18	-0.148	16.36	1,301
17	46.99 – 47.19	0.8381	108	459	2,068	0.974	100.21	-0.161	21.26	1,300
7	101.06 -101.26	0.8554	107	406	2,010	0.988	107.13	-0.157	15.85	1,297
11	71.87 -72.07	0.8593	108	403	1,975	0.990	105.71	-0.171	16.12	1,296
24	17.85 -18.05	0.8379	118	424	1,999	0.984	104.83	-0.162	21.07	1,288
3	44.53 – 44.73	0.8621	98	425	1,936	0.991	111.24	-0.154	11.04	1,284
8	64.16 – 64.36	0.8561	97	435	1,991	0.993	125.26	-0.104	5.17	1,279
13	73.89 -74.09	0.8133	113	461	2,123	0.958	107.26	-0.115	10.70	1,279
9	99.22 -99.42	0.8452	97	418	1,979	0.990	118.26	-0.122	11.20	1,276
27	26.17 -26.37	0.8482	107	439	2,065	0.986	112.71	-0.111	10.24	1,274
1	84.50 -84.70	0.8308	104	440	1,904	0.987	106.00	-0.170	22.61	1,269
11	0.16 – 0.36	0.8432	103	416	1,939	0.992	119.57	-0.125	10.28	1,265
24	20.98 – 21.18	0.8512	107	435	1,957	0.987	100.10	-0.171	19.41	1,263
15	69.69 – 69.89	0.8244	112	470	1,976	0.985	112.43	-0.123	13.27	1,250

16	30.86 -31.06	0.8017	111	508	2,148	0.958	114.04	-0.179	4.75	1,249
7	58.48 -58.68	0.8522	95	430	1,871	0.991	107.81	-0.159	13.05	1,245
23	27.37 -27.57	0.8589	97	405	1,925	0.990	104.02	-0.162	15.04	1,241
7	11.66 -11.86	0.8317	104	466	1,982	0.980	103.78	-0.143	18.01	1,239
4	8.44 -8.64	0.8438	108	435	1,945	0.986	101.16	-0.161	17.29	1,237
10	97.31 – 97.51	0.8495	95	422	1,824	0.992	113.70	-0.143	8.38	1,236
22	0.29 – 0.49	0.8380	102	448	2,017	0.988	115.86	-0.106	10.30	1,235
22	40.73 – 40.93	0.8630	101	402	1,952	0.989	102.10	-0.154	14.18	1,234
16	27.43 – 27.63	0.8024	120	505	2,134	0.915	102.31	-0.104	5.12	1,230
3	21.65 – 21.85	0.7961	109	486	2,126	0.912	102.74	-0.104	3.69	1,227

[1]Chr: Chromosome; BTA: *Bos taurus* autosome. [2]The region reported is in megabases (Mb) and corresponds to the ARS-UCD1.2 *Bos taurus* genome assembly (Rosen et al., 2020). [3]Frequency of the haplotype allele in the population. [4]Number of heterozygous sires. [5]Number of heterozygous dams. [6]Informative offspring were defined as the number of progenies derived from one or both heterozygous parents. [7]α_g, and δ_g represent the additive and dominance TRD effect, respectively. The Log_{10} (BF) indicates the Bayes Factor for the TRD effect. [8]The number of underrepresented offspring refers to the total expected offspring that were not observed for a given allele. Full list containing the 59 genomic regions harboring potential semi-lethal or lethal alleles showing TRD allelic patterns can be found in Supplementary File S12, Table S8.

Forty-four TRD genomic regions identified in the present study (Table 16) overlapped lethal haplotype regions in Nellore cattle previously reported by Schmidt et al. (2023) and Rodrigues et al. (2025). These regions were distributed across BTA1, BTA2, BTA3, BTA7, BTA8, BTA9, BTA10, BTA12, BTA13, BTA14, BTA15, BTA16, BTA17, BTA18, BTA22, BTA24, BTA25, BTA26, BTA27, and BTA29. Among the identified regions, 28 were classified as overall TRD, seven as parent-specific TRD (sire TRD), and nine as genotypic TRD. The number of underrepresented offspring in these regions ranged from 421 to 1,344. Notably, the TRD regions on BTA1 (40.84–41.04 Mb), BTA16 (1.09 – 1.29 Mb), and BTA22 (46.60–46.80 Mb) were also identified in this study as potential loci harboring semi-lethal or lethal alleles due to the presence of over 1,000 underrepresented offspring.

Table 16 Genomic regions with transmission ratio distortion (TRD) that have previously been reported as potential carriers of recessive lethal haplotypes in Nellore cattle.

Chr[1]	Region[2]	Freq[3]	TRD type	TRD effects[4]	Number of underrepresented offspring[5]	Previously reported by
BTA1	21.63 - 21.83	0.171	Genotypic	α_g = -0.98; δ_g = -0.12	920	Schmidt et al. (2023)
BTA1	36.34 - 36.54	0.945	Overall	α = 0.37	872	Rodrigues et al. (2025)
BTA1	40.84 - 41.04	0.847	Genotypic	α_g = 0.99; δ_g = -0.13	1,344	Rodrigues et al. (2025)
BTA1	81.66 - 81.89	0.949	Parent-specific	α_s = 0.36; α_d = 0.32	511	Rodrigues et al. (2025)
BTA1	92.14 - 92.34	0.094	Overall	α = -0.31	966	Schmidt et al. (2023)
BTA2	27.51 - 27.71	0.028	Overall	α = -0.41	828	Rodrigues et al. (2025)
BTA2	126.03 - 126.23	0.048	Overall	α = -0.37	754	Schmidt et al. (2023)
BTA3	84.42 - 84.62	0.944	Overall	α = 0.37	889	Schmidt et al. (2023)
BTA7	20.27 - 20.47	0.055	Overall	α = -0.37	733	Schmidt et al. (2023)
BTA7	36.96 - 37.16	0.910	Overall	α = 0.31	893	Schmidt et al. (2023) and Rodrigues et al. (2025)
BTA7	52.99 - 53.19	0.153	Genotypic	α_g = -0.99; δ_g = -0.13	835	Rodrigues et al. (2025)
BTA7	81.35 - 81.55	0.949	Parent-specific	α_s = 0.38; α_d = 0.35	871	Schmidt et al. (2023)
BTA7	106.9 - 107.1	0.061	Overall	α = -0.34	719	Schmidt et al. (2023)
BTA8	22.54 - 22.74	0.044	Parent-specific	α_s = -0.39; α_d = -0.36	818	Rodrigues et al. (2025)
BTA9	48.54 - 48.74	0.060	Overall	α = -0.36	796	Schmidt et al. (2023)
BTA9	72.77 - 72.97	0.091	Overall	α = -0.30	827	Schmidt et al. (2023)
BTA10	22.36 - 22.56	0.949	Overall	α = 0.38	831	Rodrigues et al. (2025)
BTA10	24.7 - 24.9	0.043	Overall	α = -0.40	914	Rodrigues et al. (2025)
BTA10	86.13 - 86.38	0.195	Genotypic	α_g = -0.88; δ_g = 0.01	421	Schmidt et al. (2023)
BTA12	16.33 - 16.53	0.052	Parent-specific	α_s = -0.39; α_d = -0.34	989	Rodrigues et al. (2025)
BTA13	13.33 - 13.53	0.933	Overall	α = 0.36	835	Rodrigues et al. (2025)
BTA13	15.89 - 16.09	0.190	Genotypic	α_g = -0.94; δ_g = -0.07	771	Rodrigues et al. (2025)

BTA13	16.09 - 16.29	0.960	Overall	$\alpha = 0.40$	809	Rodrigues et al. (2025)
BTA14	56.98 - 57.18	0.928	Overall	$\alpha = 0.33$	855	Rodrigues et al. (2025)
BTA15	21.13 - 21.33	0.189	Genotypic	$\alpha_g = -0.92$; $\delta_g = -0.05$	740	Rodrigues et al. (2025)
BTA15	22.96 - 23.16	0.209	Genotypic	$\alpha_g = -0.89$; $\delta_g = -0.06$	719	Rodrigues et al. (2025)
BTA16	1.09 - 1.29	0.810	Genotypic	$\alpha_g = 0.93$; $\delta_g = -0.09$	1,193	Rodrigues et al. (2025)
BTA17	67.04 - 67.24	0.056	Parent-specific	$\alpha_s = -0.39$; $\alpha_d = -0.33$	961	Rodrigues et al. (2025)
BTA17	70.36 - 70.56	0.878	Overall	$\alpha = 0.28$	913	Rodrigues et al. (2025)
BTA18	9.91 - 10.11	0.946	Parent-specific	$\alpha_s = 0.38$; $\alpha_d = 0.34$	867	Schmidt et al. (2023)
BTA18	28.3 - 28.5	0.047	Overall	$\alpha = -0.38$	791	Rodrigues et al. (2025)
BTA18	36.88 - 37.08	0.146	Genotypic	$\alpha_g = -0.99$; $\delta_g = -0.10$	828	Schmidt et al. (2023)
BTA18	54.94 - 55.14	0.882	Overall	$\alpha = 0.31$	943	Rodrigues et al. (2025)
BTA18	63.8 – 64	0.968	Parent-specific	$\alpha_s = 0.41$; $\alpha_d = 0.44$	906	Rodrigues et al. (2025)
BTA22	19.4 - 19.6	0.069	Overall	$\alpha = -0.34$	897	Schmidt et al. (2023)
BTA22	46.6 - 46.8	0.121	Overall	$\alpha = -0.29$	1,051	Schmidt et al. (2023) and Rodrigues et al. (2025)
BTA24	18.6 - 18.8	0.060	Overall	$\alpha = -0.35$	849	Schmidt et al. (2023)
BTA25	4.21 - 4.41	0.949	Overall	$\alpha = 0.38$	924	Rodrigues et al. (2025)
BTA25	12.73 - 12.93	0.095	Overall	$\alpha = -0.32$	959	Rodrigues et al. (2025)
BTA26	12.86 - 13.06	0.074	Overall	$\alpha = -0.33$	890	Rodrigues et al. (2025)
BTA26	14.89 - 15.09	0.057	Overall	$\alpha = -0.39$	882	Rodrigues et al. (2025)
BTA27	17.58 - 17.78	0.059	Overall	$\alpha = -0.35$	887	Rodrigues et al. (2025)
BTA29	40.66 - 40.86	0.058	Overall	$\alpha = -0.36$	853	Schmidt et al. (2023)
BTA29	48.41 - 48.61	0.938	Overall	$\alpha = 0.36$	849	Rodrigues et al. (2025)

[1]Chr: chromosome; BTA: *Bos taurus* autosome. [2]The region reported is in megabases (Mb) and corresponds to the ARS-UCD1.2 *Bos taurus* genome assembly (Rosen et al., 2020). [3]Frequency of the haplotype allele in the population that is under-transmitted. [4]α, α_s, α_d, α_g, and δ_g represent the overall, sire, dam, additive, and dominance TRD effect, respectively. [5]The number of underrepresented offspring refers to the total expected offspring that were not observed for a given allele (Id Lahoucine et al., 2020; Id Lahoucine et al., 2023a).

4.3.3. Genes, QTLs, and functional enrichment for the TRD regions

Considering all genomic regions (1,249) associated with TRD events, 2,265 genes were annotated. These included one long non-coding RNA (lncRNA), 90 microRNAs (miRNAs), one miscellaneous RNA (misc_RNA), 2,140 protein-coding genes, five pseudogenes, one ribosomal RNA (rRNA), 18 small nucleolar RNAs (snoRNAs), seven small nuclear RNAs (snRNAs), one T-cell receptor variable gene (*TR_V_gene*), and one Y RNA. A comprehensive list detailing the genes identified, their genome locations, and biotypes is provided in Supplementary File 13, Table S9.

Table 17 provides an overview of the QTL identified within genomic regions exhibiting TRD events. Twenty distinct enriched QTL (P-value < 0.05), associated with production (12) and reproduction (8) traits, were observed. Among the production-related QTL, traits related to growth (average daily gain, body circumference, body height, body length, metabolic body weight, maturity rate, and rump width) and feed efficiency (RFI and dry matter intake) significantly overlapped with TRD regions. Reproductive traits included QTL associated with female reproductive performance (age at first calving, calving ease, gestation length, inhibin level, and interval from first to last insemination), male reproductive traits (inhibin level and scrotal circumference), and calf survival (stillbirth). Notably, a genotypic TRD region on BTA10 (86.13–86.33 Mb) was significantly associated with calving ease (P-value < 0.0001) and overlapped with 366 QTL for this trait.

Table 17 Quantitative Trait Loci (QTL) enrichment in genomic regions showing transmission ratio distortion (TRD) in Nellore cattle

QTL Name	Number of QTLs	Chr[1]	TRD region (Mb)	TRD Type	P-value	QTL Type
Average daily gain	8	BTA29	40.66 - 40.86	Overall	0.0028	Production
Body circumference	4	BTA13	43.84 - 44.04	Overall	0.0232	Production
Body height	14	BTA8	3.33 - 3.53	Overall	<0.0001	Production
Body length	6	BTA10	91.22 - 91.42	Genotypic	0.0008	Production
Cannon bone circumference	3	BTA10	52.13 - 52.33	Overall	0.0339	Production
Dry matter intake	10	BTA2	107.55 - 107.75	Overall	0.0453	Production
Dry matter intake	5	BTA11	6.69 - 6.89	Parent-Specific (Sire)	0.0424	Production
Dry matter intake	25	BTA20	4.74 - 4.94	Parent-Specific (Sire)	<0.0001	Production
Length of productive life	13	BTA29	47.80 - 48.00	Overall	0.0002	Production
Maturity rate	9	BTA21	5.36 - 5.56	Overall	0.0008	Production
Metabolic body weight	9	BTA1	63.63 - 63.83	Overall	0.0156	Production
Residual feed intake	23	BTA14	44.53 - 44.73	Overall	<0.0001	Production
Rump width	14	BTA8	3.33 - 3.53	Overall	<0.0001	Production
Age at first calving	18	BTA2	117.36 - 117.56	Genotypic	<0.0001	Reproduction
Calving ease	17	BTA10	86.13 - 86.33	Genotypic	0.0075	Reproduction
Calving ease	349	BTA21	2.68 - 2.88	Overall	<0.0001	Reproduction
Gestation length	8	BTA29	48.41 - 48.61	Overall	0.0015	Reproduction
Inhibin level	47	BTA5	42.69 - 42.89	Overall	<0.0001	Reproduction
Interval from first to last insemination	60	BTA17	68.60 - 68.80	Overall	<0.0001	Reproduction
Scrotal circumference	40	BTA5	53.96 - 53.16	Parent-Specific (Sire)	<0.0001	Reproduction
Stillbirth	13	BTA10	86.13 - 86.33	Genotypic	0.0396	Reproduction

[1]Chr: Chromosome; BTA: *Bos taurus* autosome.

In the functional enrichment analyses of genes within TRD regions, 200 GO terms (including 129 BP, 29 CC, and 42 MF), and two KEGG pathways were significatively found after applying the Bonferroni correction for multiple testing (P-value < 0.05). The full list including all the biological processes, their significance level, and associated genes can be found in Supplementary File S14, Table S10. The results revealed expressive overrepresentation of terms related to reproductive and developmental processes. Key biological processes identified included embryo development (GO:0009790), embryonic morphogenesis (GO:0048598), animal organ morphogenesis (GO:0009887), and development of primary female sexual characteristics (GO:0046545). Processes linked to cellular and molecular dynamics, such as G2/M transition of the mitotic cell cycle (GO:0000086), cell differentiation (GO:0030154), and cell morphogenesis (GO:0000902), were also identified. Functional enrichment of terms associated with signaling and growth regulation, including response to growth factor (GO:0070848) and cell-cell signaling (GO:0007267), further highlights the importance of these genomic regions in key developmental functions. Molecular functions such as voltage-gated ion channel activity (GO:0005244) and hormone receptor activity and cellular components including synaptic membrane (GO:0097060) and plasma membrane-bounded cell projections (GO:0120039 and GO:0120036) were also observed.

4.4. Discussion

The identification of genomic regions associated with TRD events in Zebu breeds can provide new perceptions into the genetic mechanisms driving deviations from Mendelian inheritance in cattle. Despite growing interest in TRD phenomena across various livestock species, research specifically focusing on Nellore cattle remains limited. The growing availability of genotyped animals and advancements in genotyping technologies, particularly high-density SNP chips and whole-genome sequencing (WGS), have enabled the comprehensive study of TRD regions across the genome of livestock species. SNP chips allow the simultaneous genotyping of thousands of loci across the genome, being an effective tool for detecting subtle deviations in allelic transmission patterns that may indicate the presence of TRD events. Thus, this study represents an important advancement in the detection and characterization of allelic and genotypic TRD regions in Nellore cattle. By addressing

this gap, our study offers valuable perceptions into the biological basis of TRD in Nellore cattle, providing opportunities to improve reproductive outcomes and enhance the sustainability of tropical beef cattle production. The findings also have the potential to highlight novel genomic regions that may harbor semi-lethal or lethal alleles, with possible implications for biological processes and metabolic pathways that influence reproductive efficiency and genetic progress in commercial populations.

4.4.1. Characterization of TRD across the genome

In total, 37,783 SNPs and 174,190 haplotypes were associated with TRD events (Table 13). The number of haplotypes showing patterns of TRD was higher than the number of SNPs due to the use of sliding windows of varying sizes to define haplotypes. A single SNP that shows evidence of TRD can appear in multiple sliding haplotype windows, potentially leading to multiple haplotypes sharing the same TRD effect. This is a common occurrence because haplotypes are defined as groups of SNPs that are inherited together due to their proximity and LD on the chromosome, and each haplotype window may contain the same TRD-associated SNP. Thus, while the number of SNPs reflects individual loci, the higher number of haplotypes reflects the multiple overlapping locations in which those SNPs contribute to TRD patterns.

The SNPs and haplotypes identified with TRD were further grouped into genomic regions based on LD between adjacent SNPs in the same BTA. As a result, 1,249 genomic regions distributed across all autosomal chromosomes were associated with TRD events (Table 13 and Figure 6). The genetic variant with the highest BF within each genomic region exhibiting evidence of TRD was selected to represent the TRD event in that region. These regions were categorized into three types based on the nature of the transmission distortion observed: overall TRD (802 regions), parent-specific TRD (191 regions), and genotypic TRD (256 regions). Overall TRD refers to regions where a general distortion in allele transmission is observed across both parents, suggesting the presence of alleles that are preferentially transmitted to offspring regardless of parental origin (Casellas et al., 2014, 2020). Parent-Specific TRD is characterized by distortion patterns where one parent (either the sire or dam) exhibits a stronger transmission bias than the other (Casellas et al., 2014, 2020). This type of TRD can indicate the involvement of imprinted genes or parent-of-origin effects, which can have implications for understanding complex

inheritance patterns. Genotypic TRD, on the other hand, reflects variations in allele transmission depending on the genotype of the offspring, highlighting potential dominance or additive effects in genetic inheritance (Casellas et al., 2020).

The widespread distribution of TRD regions across all BTAs highlights the persistence of this phenomenon throughout the genome. In the context of the evaluated population, which is a closed herd with three selection lines subjected to distinct selection criteria for YW and RFI, the prevalence of TRD regions may be attributed to the cumulative effects of intense selection pressures. This could have driven deviations from Mendelian inheritance in loci associated with the targeted traits, particularly in genomic regions where alleles conferred either selective advantages or disadvantages. Furthermore, the presence of TRD regions may reflect biological processes influenced by deleterious alleles, such as embryonic lethality, reduced fertility, or impaired development, which can alter transmission patterns and disrupt normal inheritance (VanRaden et al., 2011; Schmidt et al., 2023). Additionally, the identification of TRD regions provides valuable insights into the genetic variation that may not be captured by traditional association studies, offering opportunities to refine genomic selection strategies (Id-Lahoucine et al., 2023a; Id-Lahoucine et al., 2023b).

The variation in TRD coefficients (α values) across the genome highlights the complexity of the underlying inheritance mechanisms. A diverse array of TRD patterns was observed, including regions with consistent transmission bias and others exhibiting contrasting transmission depending on the direction of inheritance. These patterns suggest the involvement of non-Mendelian processes, such as genomic imprinting, epistatic interactions, or other parent-of-origin effects that can influence allele segregation (Huang et al., 2013; Fishman and Mcintosh, 2019). Notably, one genomic region showed clear divergence in transmission direction depending on the parental origin of the allele, indicating potential interactions between distinct regulatory pathways during gametogenesis or early embryonic development. These effects may result from allele-specific interactions, disruptions in zygotic compatibility, or coordinated epigenetic modifications (Id-Lahoucine et al., 2019; 2023a; 2023c). Further investigation is necessary to determine whether such patterns reflect a unified TRD signal or distinct mechanisms acting concurrently in the same region. While sex-specific differences may be partially explained by sample size or LD, the consistency of directionally opposing transmission implies that additional biological factors likely contribute to these patterns.

The predominance of sire-specific TRD regions (183) compared to dam-specific regions (7) suggests a potential biological basis for this disparity. One contributing factor may be that males produce significantly more sperm cells than females produce ova (Siu et al., 2021). Furthermore, sperm cells undergo more rounds of cell division during spermatogenesis than oocytes do during oogenesis (Siu et al., 2021), increasing the likelihood of mutations that could distort allele transmission patterns. Understanding the nature of these TRD effects is crucial for developing more targeted genetic selection approaches, particularly in the context of traits influenced by complex genetic interactions and parent-of-origin effects (Huang et al., 2013; Casellas et al., 2017; Fishman and Mcintosh, 2019).

The genotypic parametrization of TRD revealed the relative contributions of additive and dominance effects to transmission patterns. Notably, additive genetic effects emerged as the prevalent factor influencing genotypic TRD events, as evidenced by higher BF values for additive effects compared to dominance effects. This emphasizes the central role of additive variation, which reflects the cumulative contribution of individual alleles, in driving TRD and shaping allele transmission patterns. In contrast, dominance effects, resulting from interactions between alleles at a locus, played a comparatively minor role in these regions. These findings have important implications for breeding programs, as additive effects are more straightforward to estimate and are consistently transmitted to subsequent generations. Due to their direct influence and predictability, additive genetic effects are included in selection indexes to enhance traits through animal breeding (Falconer and Mackay, 1996). Recognizing the predominance of additive effects in genotypic TRD events can be useful for refining breeding strategies, ultimately improving the efficiency of genomic selection in traits regulated by complex inheritance mechanisms.

4.4.2. Identification of potential candidates for semi-lethal or lethal alleles in TRD regions

The identification of TRD regions harboring potential semi-lethal or lethal alleles represents opportunities for livestock breeding. From a population genetics standpoint, these alleles are likely subject to purifying selection, which would normally reduce their frequency over time in natural populations (Cvijović et al., 2018; Buffalo and Kern, 2024; Hasan and Whitlock, 2024). However, in the context of animal

breeding programs, lethal alleles could persist or even accumulate within subpopulations if not carefully managed (VanRaden et al., 2011; Schmidt et al., 2023; Rodrigues et al., 2025). Therefore, incorporating TRD analyses into breeding strategies is crucial for the effective management of genomic regions that may impact the long-term viability and sustainability of beef cattle production.

By employing criteria to select allelic TRD genomic regions potentially harboring semi-lethal or lethal alleles - defined as regions with a minimum of 1,000 underrepresented offspring and a TRD effect magnitude of |0.20|, a total of 73 genomic regions were identified (Table 14 and Supplementary File S2, Table S2). Of these, 45 were associated with overall TRD, while 28 were linked to parent-specific TRD (27 sire-specific and 1 dam-specific), highlighting the diversity of the mechanisms driving the TRD phenomena. The identification of these genomic regions underscores their potential consequences for animal breeding, as they may harbor alleles that disrupt normal transmission patterns, leading to reduced reproductive success or the perpetuation of harmful genetic variants. Understanding these mechanisms is crucial for refining breeding programs, as it allows for the strategic management of alleles that can significantly impact livestock efficiency and sustainability. The TRD region on BTA23 (32.46–32.66 Mb) exhibited 1,092 underrepresented offspring (α_s = -0.379 ± 0.010), suggesting its association with a lethal or semi-lethal allele. Furthermore, this region has also been associated with pregnancy loss in Nellore cattle based on a genome-wide association study (GWAS) (Rodrigues et al., 2024b). Similarly, the TRD region on BTA19 (59.76–59.96 Mb), which exhibited 1,114 underrepresented offspring (α = 0.373 ± 0.008), was also identified in a GWAS for stillbirth in Nellore cattle (Rodrigues et al., 2024b).

The genotypic parametrization of TRD also revealed 59 genomic regions potentially harboring semi-lethal or lethal alleles (Table 15 and Supplementary File S3, Table S3). These regions showed strong evidence of additive TRD effects and in some of these regions, heterozygous-by-heterozygous matings (AB × AB) resulted in the complete absence of homozygous offspring (AA or BB), exhibiting signs of lethal outcomes. For instance, genomic regions on BTA1, BTA8, BTA13, BTA24, and BTA27 exhibited this pattern, suggesting the presence of alleles that impair viability when homozygous. This finding emphasizes the potential biological consequences of TRD regions, as such alleles may disrupt key developmental or physiological pathways, leading to embryonic lethality or severe fitness reductions. The genotypic TRD region

on BTA18 (64.52–64.72 Mb) was also detected in a GWAS for pre-weaning mortality in Nellore cattle (Rodrigues et al., 2024b).

The TRD analyses conducted in this study identified 44 genomic regions that overlap with the findings of Schmidt et al. (2023) and Rodrigues et al. (2025) (Table 16), both of which examined genomic regions as potential candidates harboring lethal haplotypes based on the absence of homozygous carriers of recessive haplotypes. These findings emphasize the consistency of TRD as a marker for identifying genomic regions that may influence survival and reproduction. Furthermore, these genomic regions are potential hotspots for lethal or semi-lethal alleles and may have potential mutations in specific genes associated with early mortality of specific genotypes in Nellore cattle. The identification of these common regions across independent studies provides strong groundwork for subsequent investigations using WGS, gene expression, and/or transcriptomic analyses to comprehend the genetic mechanisms underlying lethal alleles. Notably, genomic regions on BTA1 and BTA22 were consistently identified across this study as containing lethal alleles, and in the study of Schmidt et al. (2023) and Rodrigues et al. (2025) as candidate regions harboring lethal haplotypes, suggesting their pivotal role in reproductive fitness and embryonic viability in Nellore cattle.

From an animal breeding perspective, monitoring TRD regions associated with semi-lethal or lethal alleles is essential to avoid the spread of deleterious alleles, which could reduce genetic gains and compromise long-term herd efficiency. By identifying TRD regions it is possible to design mating strategies that minimize the transmission of deleterious alleles. For instance, avoiding matings that are likely to produce homozygous recessive offspring can reduce embryonic lethality and improve reproductive efficiency. Future research should focus on the functional characterization of these TRD regions to elucidate the specific biological mechanisms by which deleterious alleles influence fitness, thereby enabling more effective management strategies in cattle breeding programs. By accounting for TRD in breeding programs, it may be possible to enhance genetic progress while mitigating undesirable outcomes, such as the inadvertent propagation of deleterious alleles. This approach could lead to more sustainable genetic improvement in commercial populations.

4.4.3 Biological aspects of TRD regions: genes, QTLs, and functional enrichment

The annotation of 2,265 genes overlapping with TRD regions reveals a diverse array of gene biotypes, each potentially contributing to the transmission bias observed in these loci (Supplementary file 4 Table S4). Protein-coding genes were predominant in TRD regions, reflecting their likely role in fundamental biological pathways such as metabolism, structural integrity, and regulatory functions. These genes are often key players in processes influencing key traits in cattle (Brunes et al., 2021; Mota et al., 2022; Rodrigues et al., 2024b; Souza et al., 2024). Regulatory non-coding RNAs, including 90 miRNAs, 18 snoRNAs, and one lncRNA, underscore the complexity of gene regulation within TRD regions. Non-coding RNAs regulate gene expression through transcriptional repression, RNA splicing, and post-transcriptional modification (Kaikkonen et al., 2011; Zhao et al., 2016). These elements could mediate the genetic imbalance aspect of TRD by modulating key developmental or reproductive processes. For instance, miRNAs are known to influence developmental timing and stress responses, roles that are critical during early embryogenesis—a period particularly sensitive to genetic imbalances (Mendell and Olson, 2012; Marsico et al., 2023). Additional biotypes, such as rRNAs, pseudogenes, and Y RNAs, emphasize the involvement of fundamental cellular processes in TRD. Ribosome biogenesis and RNA modification, driven by snoRNAs and rRNAs, are essential for cellular growth and function (Sloan et al., 2017; Ojha et al., 2020), linking these elements to the observed enrichment of QTLs for production traits. These QTLs include growth (e.g., average daily gain, metabolic body weight, and maturity rate) and feed efficiency (e.g., RFI and dry matter intake), suggesting that TRD regions may influence phenotypes associated with energy utilization and developmental efficiency. Such traits are under strong selection pressure in cattle breeding programs, making TRD regions key targets for further exploration.

Reproductive traits also showed significant associations with TRD regions, underscoring their importance to fertility and survival. QTLs linked to female reproductive performance, such as age at first calving, calving ease, and gestation length, and male reproductive traits, such as scrotal circumference, emphasize the genetic mechanisms underlying reproductive success. The TRD region on BTA10 (86.13–86.33 Mb) shows this connection, indicating a significant association with calving ease and overlapping with 366 QTLs for this trait. This region may harbor alleles with pleiotropic effects, influencing maternal and neonatal viability during

calving through variations in uterine environment, fetal growth, or pelvic structure. Associations between TRD regions and QTLs for stillbirth further suggest a role in calf survival. These findings support the hypothesis that some TRD loci harbor alleles with deleterious effects, which are maintained in populations through balancing selection or heterozygote advantage (Hedrick, 2012; Derks and Steensma, 2021). For example, alleles that impair early development when homozygous but confer advantages in heterozygous individuals could persist due to their net fitness benefit (Hedrick, 2012; Derks and Steensma, 2021).

The functional enrichment analyses of genes within TRD regions highlights their potential roles in fertility, early embryogenesis, and reproductive success. Reproductive terms such as embryo development and embryonic morphogenesis underscore the importance of genes involved in early developmental stages for producing viable offspring. The TRD phenomena may arise from alleles that either enhance or disrupt these critical processes, leading to selection-driven distortions in allele transmission. The presence of enriched terms such as development of primary female sexual characteristics suggests that TRD regions harbor genes crucial for the formation and function of reproductive traits. These genes likely regulate processes such as ovarian folliculogenesis, hormonal signaling pathways, and uterine receptivity, which are functions essential for reproductive success (Christensen et al., 2012; Fiorentino et al., 2023). Mutations or variations within these regions might lead to reduced fertility or embryonic lethality, particularly when deleterious alleles are homozygous. Processes related to cell differentiation and cell morphogenesis play key roles in early embryonic development. These mechanisms control the specialization and organization of cells into functional tissues and organs, which are essential for zygotic development and successful implantation (Huang et al., 2025). Mutations in these processes may delay the development of embryo, potentially leading to embryonic loss (Das and Holzer, 2012; Huang et al., 2025). Ion channel activity, particularly voltage-gated ion channel activity, appears to be another important process enriched in TRD regions. Ion channels regulate key events such as oocyte maturation, sperm motility, and calcium oscillations during oocyte activation and early embryogenesis (Darszon et al., 2011; Xu and Yang, 2017). Dysfunctions in ion transport mechanisms caused by mutations in these loci could impair reproductive processes and contribute to allele transmission distortions.

4.4.4. Future directions for research

The identification of TRD regions within the Nellore cattle genome, particularly those harboring lethal or semi-lethal alleles, opens up significant possibilities for advancing animal breeding strategies. The incorporation of TRD regions into genomic selection models represents an important advance in animal breeding. Genomic selection, which has already revolutionized breeding programs by selecting animals based on their genomic information, can be further optimized using TRD information. By monitoring the transmission of alleles, breeders can make more informed decisions, avoiding the perpetuation of lethal alleles that may result in reduced fertility or embryonic mortality. Moreover, while this study has identified several candidate genes within TRD regions, the functional characterization of nearly all these genes in Nellore cattle is still in its early stages, especially regarding their role in reproductive mechanisms. Future research should focus on experimentally validating these candidate genes, using techniques such as transcriptomics and proteomics, to gain a deeper understanding of their biological functions and their contribution to generating TRD patterns.

4.5. Conclusions

We identified 1,249 TRD regions distributed across the Nellore cattle genome. Of these, 73 and 59 TRD regions were classified as potential genomic hotspots harboring semi-lethal or lethal alleles, identified through allelic and genotypic parametrizations of TRD effects, respectively. Enrichment analyses revealed that TRD regions are strongly associated with key QTLs, including growth, feed efficiency, and reproductive performance, underscoring their importance in livestock productive efficiency. Functional annotation of genes within TRD regions highlighted significant biological processes, such as early embryogenesis, gametogenesis, cellular signaling, and ion transport. These findings emphasize the importance of incorporating TRD information into genomic selection strategies to enhance productivity while minimizing the segregation of deleterious alleles in cattle populations.

4.6. Supplementary files

All supplementary materials have been deposited in the Harvard Dataverse repository and are available at the permanent link: https://doi.org/10.7910/DVN/T9HCLU

4.7. Disclosures

The authors declare that they have no competing interests.

4.7.1. Data availability

The raw data can be required by contacting MEZM (e-mail: mezmercadante@gmail.com) upon a reasonable request for research purposes.

4.7.2. Acknowledgements

The authors thank the staff of the Institute of Animal Science - Beef Cattle Research Center (Sertãozinho, SP, Brazil) for their efforts in collecting the data used in this study, and São Paulo Research Foundation (FAPESP) for financial support.

4.7.3 Funding information

The São Paulo Research Foundation (FAPESP) provided financial support (2017/10630-2 and 2017/50339-5) and scholarships to GRDR (2023/11176-4 and 2024/01909-7) and LFMM (2022/11852-7). The National Council for Science and Technological Development (CNPq) provided financial support through competitive grants to LEF, LGA, and MEZM.

References

Abdalla, E. A., S. Id-Lahoucine, A. Cánovas, J. Casellas, F. S. Schenkel, B. J. Wood, and C. F. Baes. 2020. Discovering lethal alleles across the turkey genome using a transmission ratio distortion approach. Anim Genet. 51:876–889. doi:10.1111/age.13003.

Amorim, C. E. G., Z. Gao, Z. Baker, J. F. Diesel, Y. B. Simons, I. S. Haque, J. Pickrell, and M. Przeworski. 2017. The population genetics of human disease: The case of recessive, lethal mutations. PLoS Genet. 13:e1007499. doi:10.1371/journal.pgen.1006915.

Bem, R. D., L. F. Benfica, D. A. Silva, E. R. Carrara, L. F. Brito, H. A. Mulim, M. S. Borges, J. N. S. G. Cyrillo, R. C. Canesin, S. F. M. Bonilha, and M. E. Z. Mercadante. 2024. Assessing different metrics of pedigree and genomic inbreeding and inbreeding effect on growth, fertility, and feed efficiency traits in a closed-herd Nellore cattle population. BMC Genomics. 25:738. doi:10.1186/s12864-024-10641-3.

Benfica, L. F., L. F. Brito, R. D. do Bem, H. A. Mulim, J. Glessner, L. G. Braga, L. S. Gloria, J. N. S. G. Cyrillo, S. F. M. Bonilha, and M. E. Z. Mercadante. 2024. Genome-wide association study between copy number variation and feeding behavior, feed efficiency, and growth traits in Nellore cattle. BMC Genomics. 25. doi:10.1186/s12864-024-09976-8.

Bosse, M., H. J. Megens, M. F. L. Derks, Á. M. R. de Cara, and M. A. M. Groenen. 2019. Deleterious alleles in the context of domestication, inbreeding, and selection. Evol Appl. 12:6–17. doi:10.1111/eva.12691.

Brunes, L. C., F. Baldi, F. B. Lopes, R. B. Lôbo, R. Espigolan, M. F. O. Costa, N. B. Stafuzza, and C. U. Magnabosco. 2021. Weighted single-step genome-wide association study and pathway analyses for feed efficiency traits in Nellore cattle. J Anim Breed Genet. 138:23–44. doi:10.1111/jbg.12496.

Buffalo, V., and A. D. Kern. 2024. A quantitative genetic model of background selection in humans. PLoS Genet. 20:e1011144. doi:10.1371/journal.pgen.1011144.

Casellas, J., J. J. Cañas-Álvarez, A. González-Rodríguez, A. Puig-Oliveras, M. Fina, J. Piedrafita, A. Molina, C. Díaz, J. A. Baró, and L. Varona. 2017. Bayesian analysis of parent-specific transmission ratio distortion in seven Spanish beef cattle breeds. Anim Genet. 48:93–96. doi:10.1111/age.12509.

Casellas, J., R. J. Gularte, C. R. Farber, L. Varona, M. Mehrabian, E. E. Schadt, A. J. Lusis, A. D. Attie, B. S. Yandell, and J. F. Medrano. 2012. Genome scans for transmission ratio distortion regions in mice. Genetics. 191:247–259. doi:10.1534/genetics.111.135988.

Casellas, J., S. Id-Lahoucine, and A. Cánovas. 2020. Discriminating between allele- and genotype-specific transmission ratio distortion. Anim Genet. 51:847–854. doi:10.1111/age.13007.

Casellas, J., A. Manunza, A. Mercader, R. Quintanilla, and M. Amills. 2014. A flexible bayesian model for testing for transmission ratio distortion. Genetics. 198:1357–1367. doi:10.1534/genetics.114.169607.

Christensen, A., G. E. Bentley, R. Cabrera, H. H. Ortega, N. Perfito, T. J. Wu, and P. Micevych. 2012. Hormonal regulation of female reproduction. Hormone and Metabolic Research. 44:587–591. doi:10.1055/s-0032-1306301.

Cvijović, I., B. H. Good, and M. M. Desai. 2018. The effect of strong purifying selection on genetic diversity. Genetics. 209:1235–1278. doi:10.1534/genetics.118.301058.

Cyrillo, N. dos S. G. J., A. G. Razook, L. Andrade De Figueiredo, L. Martins, B. Neto, M. Eugênia, Z. Mercadante, and H. Tonhati. 2001. Estimates of Genetic Parameters and Trends for Body Weight at 378 Days, Body Measurements and Scrotal Circumference of Sertãozinho Nellore Males. Rev Bras Zootec. 30:56–65. doi:10.1590/S1516-35982001000100010

Darszon, A., T. Nishigaki, C. Beltran, and C. L. Treviño. 2011. Calcium channels in the development, maturation, and function of spermatozoa. Physiol Rev. 91:1305–1355. doi:10.1152/physrev.00028.2010.-A.

Das, M., and H. E. G. Holzer. 2012. Recurrent implantation failure: Gamete and embryo factors. Fertil Steril. 97:1021–1027. doi:10.1016/j.fertnstert.2012.02.029.

Derks, M. F. L., and M. Steensma. 2021. Review: Balancing Selection for Deleterious Alleles in Livestock. Front Genet. 12:761728. doi:10.3389/fgene.2021.761728.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to Quantitative Genetics. 3rd ed. Pearson, New York, USA.

Fiorentino, G., D. Cimadomo, F. Innocenti, D. Soscia, A. Vaiarelli, F. M. Ubaldi, G. Gennarelli, S. Garagna, L. Rienzi, and M. Zuccotti. 2023. Biomechanical forces and signals operating in the ovary during folliculogenesis and their dysregulation: implications for fertility. Hum Reprod Update. 29:1–23. doi:10.1093/humupd/dmac031.

Fishman, L., and M. Mcintosh. 2019. Standard Deviations: The biological bases of transmission ratio distortion. Annu Rev Genet. 53:347–372. doi:10.1146/annurev-genet-112618.

Fonseca, P. A. S., A. Suárez-Vega, G. Marras, and Á. Cánovas. 2020. GALLO: An R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. Gigascience. 9:giaa149. doi:10.1093/gigascience/giaa149.

Guo, Y., J. Li, A. J. Bonham, Y. Wang, and H. Deng. 2009. Gains in power for exhaustive analyses of haplotypes using variable-sized sliding window strategy: A comparison of association-mapping strategies. Eur J Hum Genet. 17:785–792. doi:10.1038/ejhg.2008.244.

Hasan, A., and M. C. Whitlock. 2024. FST and genetic diversity in an island model with background selection. B. Payseur, editor. PLoS Genet. 20:e1011225. doi:10.1371/journal.pgen.1011225.

Hedrick, P. W. 2012. What is the evidence for heterozygote advantage selection? Trends Ecol Evol. 27:698–704. doi:10.1016/j.tree.2012.08.012.

Hu, Z. L., C. A. Park, and J. M. Reecy. 2022. Bringing the Animal QTLdb and CorrDB into the future: Meeting new challenges and providing updated services. Nucleic Acids Res. 50:956–961. doi:10.1093/nar/gkab1116.

Huang, H., S. Gao, and M. Bao. 2025. Exploring Mechanical Forces Shaping Self-Organization and Morphogenesis During Early Embryo Development. Annu Rev Cell Dev Biol. 106:4. doi:10.1146/annurev-cellbio-120123.

Huang, L. O., A. Labbe, and C. Infante-Rivard. 2013. Transmission ratio distortion: Review of concept and implications for genetic association studies. Hum Genet. 132:245–263. doi:10.1007/s00439-012-1257-0.

Id-Lahoucine, S., A. Cánovas, C. Jaton, F. Miglior, P. A. S. Fonseca, M. Sargolzaei, S. Miller, F. S. Schenkel, J. F. Medrano, and J. Casellas. 2019. Implementation of Bayesian methods to identify SNP and haplotype regions with transmission ratio distortion across the whole genome: TRDscan v.1.0. J Dairy Sci. 102:3175–3188. doi:10.3168/jds.2018-15296.

Id-Lahoucine, S., J. Casellas, A. Suárez-Vega, P. A. S. Fonseca, F. S. Schenkel, M. Sargolzaei, and A. Cánovas. 2023a. Unravelling transmission ratio distortion across the bovine genome: identification of candidate regions for reproduction defects. BMC Genomics. 24:383. doi:10.1186/s12864-023-09455-6.

Id-Lahoucine, S., A. Cánovas, A. Legarra, and J. Casellas. 2023b. Transmission ratio distortion regions in the context of genomic evaluation and their effects on reproductive traits in cattle. J Dairy Sci. 106:7786–7798. doi:10.3168/jds.2022-23062.

Id-Lahoucine, S., J. Casellas, D. Lu, M. Sargolzaei, S. Miller, and A. Cánovas. 2023c. Distortion of Mendelian segregation across the Angus cattle genome uncovering regions affecting reproduction. Sci Rep. 13:13393. doi:10.1038/s41598-023-37710-z.

Id-Lahoucine, S., J. Casellas, F. Miglior, F. S. Schenkel, and A. Cánovas. 2023d. Parent-offspring genotyped trios unravelling genomic regions with gametic and genotypic epistatic transmission bias on the cattle genome. Front Genet. 14:1-14. Doi: 10.3389/fgene.2023.1132796.

Kaikkonen, M. U., M. T. Y. Lam, and C. K. Glass. 2011. Non-coding RNAs as regulators of gene expression and epigenetics. Cardiovasc Res. 90:430–440. doi:10.1093/cvr/cvr097.

Kass, R. E., and A. E. Raftery. 1995. Bayes Factors. Journal of American Statistical Association. 90:773–795.

Kolberg, L., U. Raudvere, I. Kuzmin, P. Adler, J. Vilo, and H. Peterson. 2023. G:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res. 51:207–212. doi:10.1093/nar/gkad347.

Laseca, N., Á. Cánovas, M. Valera, S. Id-Lahoucine, D. I. Perdomo-González, P. A. S. Fonseca, S. Demyda-Peyrás, and A. Molina. 2023. Genomic screening of allelic and genotypic transmission ratio distortion in horse. PLoS One. 18:e0289066. doi:10.1371/journal.pone.0289066.

Li, Y., W. K. Sung, and J. J. Liu. 2007. Association mapping via regularized regression analysis of single-nucleotide-polymorphism haplotypes in variable-sized sliding windows. Am J Hum Genet. 80:705–715. doi:10.1086/513205.

Marsico, T. V., M. V. Silva, R. S. Valente, K. Annes, V. B. Rissi, W. G. Glanzner, and M. J. Sudano. 2023. Unraveling the Consequences of Oxygen Imbalance on Early Embryo Development: Exploring Mitigation Strategies. Animals. 13:1–19. doi:10.3390/ani13132171.

Mendell, J. T., and E. N. Olson. 2012. MicroRNAs in stress signaling and human disease. Cell. 148:1172–1187. doi:10.1016/j.cell.2012.02.005.

Mercadante, M. E. Z., I. U. Packer, A. G. Razook, J. N. S. G. Cyrillo, and L. A. Figueiredo. 2003. Direct and correlated responses to selection for yearling weight on reproductive performance of Nelore cows. J Anim Sci. 81:376–384. doi:10.2527/2003.812376x

Mota, L. F. M., F. B. Lopes, G. A. Fernandes Júnior, G. J. M. Rosa, A. F. B. Magalhães, R. Carvalheiro, and L. G. Albuquerque. 2020. Genome-wide scan highlights the role of candidate genes on phenotypic plasticity for age at first calving in Nellore heifers. Sci Rep. 10:6481. doi:10.1038/s41598-020-63516-4.

Mota, L.F.M., S. W. B. Santos, G. A. F. Júnior, T. Bresolin, M. E. Z. Mercadante, J. A. V Silva, J. N. S. G. Cyrillo, F. M. Monteiro, R. Carvalheiro, and L. G. Albuquerque. 2022. Meta-analysis across Nellore cattle populations identifies common metabolic mechanisms that regulate feed efficiency-related traits. BMC Genomics. 23:424. doi:10.1186/s12864-022-08671-w.

Mota, L. F. M., L. M. Arikawa, S. W. B. Santos, G. A. Fernandes Júnior, A. A. C. Alves, G. J. M. Rosa, M. E. Z. Mercadante, J. N. S. G. Cyrillo, R. Carvalheiro, and L. G. Albuquerque. 2024. Benchmarking machine learning and parametric methods for genomic prediction of feed efficiency-related traits in Nellore cattle. Sci Rep. 14:6404. doi:10.1038/s41598-024-57234-4.

Ojha, S., S. Malla, and S. M. Lyons. 2020. SnoRNPs: Functions in ribosome biogenesis. Biomolecules. 10:1–20. doi:10.3390/biom10050783.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. De Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 81:559–575. doi:10.1086/519795.

Rodrigues, G. R. D., V. T. Rezende, M. E. Z. Mercadante, S. F. M. Bonilha, R. C. Canesin, C. Raineri, J. D. P. S. Valente, V. A. Ligori, and J. N. D. S. Gonçalves Cyrillo. 2024a. Animal growth models as a tool to estimate resilience indicators in Bos indicus and Bos taurus heifers: Selection effects and genetics parameters. Livest Sci. 282:105435. doi:10.1016/j.livsci.2024.105435.

Rodrigues, G. R.D., L. F. Brito, L. F. M. Mota, J. N. S. G. Cyrillo, J. P. S. Valente, L. F. Benfica, J. B. Silva Neto, M. S. Borges, F. M. Monteiro, L. El Faro, L. G. Albuquerque, and M. E. Z. Mercadante. 2024b. Genome-wide association studies and functional annotation of pre-weaning calf mortality and reproductive

traits in Nellore cattle from experimental selection lines. BMC Genomics. 25:1196. doi:10.1186/s12864-024-11113-4.

Rodrigues, G. R. D., J. N. S. G. Cyrillo, P. I. Schmidt, J. P. S. Valente, E. S. Oliveira, L. G. Albuquerque, L. F. Brito, and M. E. Z. Mercadante. 2025. Effect of genomic regions harboring putative lethal haplotypes on reproductive performance in closed experimental selection lines of Nellore cattle. Sci Rep. 15:4113. doi:10.1038/s41598-025-88501-7.

Rosen, B. D., D. M. Bickhart, R. D. Schnabel, S. Koren, C. G. Elsik, E. Tseng, T. N. Rowan, W. Y. Low, A. Zimin, C. Couldrey, R. Hall, W. Li, A. Rhie, J. Ghurye, S. D. McKay, F. Thibaud-Nissen, J. Hoffman, B. M. Murdoch, W. M. Snelling, T. G. McDaneld, J. A. Hammond, J. C. Schwartz, W. Nandolo, D. E. Hagen, C. Dreischer, S. J. Schultheiss, S. G. Schroeder, A. M. Phillippy, J. B. Cole, C. P. Van Tassell, G. Liu, T. P. L. Smith, and J. F. Medrano. 2020. De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience. 9:1–9. doi:10.1093/gigascience/giaa021.

Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics. 15:1–12. doi:10.1186/1471-2164-15-478.

Schmidt, P. I., L. F. M. Mota, L. F. S. Fonseca, D. B. dos Santos Silva, G. B. Frezarim, L. M. Arikawa, D. J. de Abreu Santos, A. F. B. Magalhães, J. B. Cole, R. Carvalheiro, H. N. de Oliveira, D. J. Null, P. VanRaden, L. Ma, and L. G. de Albuquerque. 2023. Identification of candidate lethal haplotypes and genomic association with post-natal mortality and reproductive traits in Nellore cattle. Sci Rep. 13:1–12. doi:10.1038/s41598-023-37586-z.

Silva Neto, J. B., L. F. M. Mota, S. T. Amorim, E. Peripolli, L. F. Brito, C. U. Magnabosco, and F. Baldi. 2023. Genotype-by-environment interactions for feed efficiency traits in Nellore cattle based on bi-trait reaction norm models. Genetics Selection Evolution. 55:93. doi:10.1186/s12711-023-00867-2.

Siu, K. K., V. H. B. Serrão, A. Ziyyat, and J. E. Lee. 2021. The cell biology of fertilization: Gamete attachment and fusion. Journal of Cell Biology. 220:e202102146. doi:10.1083/jcb.202102146.

Sloan, K. E., A. S. Warda, S. Sharma, K. D. Entian, D. L. J. Lafontaine, and M. T. Bohnsack. 2017. Tuning the ribosome: The influence of rRNA modification on

eukaryotic ribosome biogenesis and function. RNA Biol. 14:1138–1152. doi:10.1080/15476286.2016.1259781.

Souza, L. L., P. Dominguez-Castaño, S. B. Gianvecchio, L. S. Sakamoto, G. R. D. Rodrigues, T. L. da Silva Soares, S. F. M. Bonilha, J. de Oliveira Santos Marcatto, L. G. Albuquerque, J. A. Vasconcelos Silva, and M. E. Z. Mercadante. 2024. Heritability estimates and genome-wide association study of methane emission traits in Nellore cattle. J Anim Sci. 102:skae182. doi:10.1093/jas/skae182.

VanRaden, P. M., K. M. Olson, D. J. Null, and J. L. Hutchison. 2011. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. J Dairy Sci. 94:6153–6161. doi:10.3168/jds.2011-4624.

Vázquez-Gómez, M., M. M. de Hijas-Villalba, L. Varona, N. Ibañez-Escriche, J. P. Rosas, S. Negro, J. L. Noguera, and J. Casellas. 2020. Maternal transmission ratio distortion in two iberian pig varieties. Genes (Basel). 11:1–16. doi:10.3390/genes11091050.

Wright, S. D. R. 1922. Coefficients of Inbreeding and Relationship. Am Nat. 56:330–338. doi:10.1086/279872

Xu, Y.-R., and W.-X. Yang. 2017. Calcium influx and sperm-evoked calcium responses during oocyte maturation and egg activation. Oncotarget. 8:89375–89390.

Zhao, B. S., I. A. Roundtree, and C. He. 2016. Post-transcriptional gene regulation by mRNA modifications. Nat Rev Mol Cell Biol. 18:31–42. doi:10.1038/nrm.2016.132.

CHAPTER 5: FINAL CONSIDERATIONS, PRACTICAL APPLICATIONS, AND FUTURE DIRECTIONS FOR RESEARCH

The present dissertation provided an integrated genomic perspective on reproductive performance and pre-weaning calf mortality in Nellore cattle, combining three complementary approaches: GWAS with functional annotation, the detection of putative lethal haplotypes, and the characterization of genomic regions exhibiting TRD. Together, these studies advanced current knowledge about the genetic architecture of reproduction in *Bos taurus indicus* populations, highlighting both the polygenic background of fertility-related traits and the specific contribution of deleterious variants that compromise conception, embryonic development, and early calf survival. In doing so, this study addressed important gaps in the literature on Nellore cattle and generated information with direct relevance for breeding programs that rely increasingly on genomic selection.

In the first study, GWAS were used to investigate conception success, pregnancy loss, stillbirth, and pre-weaning calf mortality in Nellore cattle. These traits are notoriously difficult to improve through traditional selection because they tend to exhibit low heritability, late expression, and strong environmental sensitivity. The results confirmed their highly polygenic nature, with some genomic regions explaining moderate proportions of additive genetic variance but also revealed specific loci and biological pathways with plausible roles in the regulation of reproductive function, gestational maintenance, and neonatal survival. Genes and pathways associated with immune response, cell cycle regulation, hormonal signaling, energy metabolism, and embryonic development were recurrently implicated, reinforcing the view that reproductive efficiency is governed by a complex interplay between endocrine, metabolic, and immunological mechanisms. The use of functional annotation and enrichment analyses was essential to move beyond a purely statistical description of associated markers and to place the GWAS signals into a coherent biological context.

The second study focused on the identification of putative lethal haplotypes in closed experimental selection lines of Nellore cattle. By exploring the distribution of haplotypes across the genome and searching for recessive segments with complete absence of homozygous carriers, it was possible to pinpoint 45 genomic regions that likely harbor recessive lethal alleles. These regions overlapped genes and pathways

associated with sexual development, gametogenesis, implantation, placental function, immune regulation, and energy homeostasis, indicating that the underlying deleterious variants may impair key stages ranging from fertilization and early embryogenesis to fetal development and perinatal survival. Importantly, the analyses of reproductive records revealed clear phenotypic consequences of these haplotypes: matings between heterozygous carriers were associated with reductions in pregnancy success and increases in pregnancy loss, stillbirth, and pre-weaning mortality. Thus, the absence of homozygotes was translated into measurable reductions in reproductive performance at the herd level. Beyond identifying novel candidates, the lethal haplotype study also converged with previous efforts in Nellore cattle, which had reported a set of homozygous-deficient haplotypes in independent populations. The recurrence of specific genomic segments across distinct datasets and population structures strongly suggests that some of these regions harbor recessive lethals of broad relevance for *Bos taurus indicus*. In practical terms, this convergence provides more robust targets for future fine-mapping studies and for the implementation of genomic tools aimed at routine carrier detection and informed mating plans in commercial herds.

The third study extended the investigation of deleterious variation by performing the first comprehensive TRD scan in Nellore cattle. Instead of focusing only on the absence of homozygous haplotypes, this analysis evaluated deviations from Mendelian expectations at both allelic and genotypic levels across the genome, considering overall TRD as well as parent-specific patterns. In total, 1,249 TRD regions were identified, including 802 regions with global allelic or genotypic distortion, 191 with parent-specific TRD, and 256 with evidence of genotypic TRD reflecting deficits of particular genotypes. Among these, 132 regions stood out as plausible carriers of semi-lethal or lethal alleles, based on the magnitude of underrepresented offspring classes and, in some cases, complete absence of specific homozygous genotypes. The functional characterization of TRD regions revealed an enrichment of genes involved in embryo development, morphogenesis, organogenesis, cell survival, and reproductive processes, as well as overlaps with known QTL associated with fertility and production traits. These findings corroborate the interpretation that at least part of the observed distortion arises from viability selection acting before birth, through early embryonic loss or fetal mortality. At the same time, the identification of parent-specific TRD patterns suggests that gametic selection, meiotic drive, or parent-

of-origin effects may also contribute to shaping allele transmission in Nellore cattle. Thus, TRD analyses provided not only an additional avenue for detecting candidate deleterious variants, but also new insights into the non-Mendelian processes that can influence genomic responses to selection.

Taken together, the three studies presented in this dissertation portray reproductive performance in Nellore cattle as the result of both a diffuse polygenic background and the discrete effects of deleterious alleles that perturb key biological processes. The GWAS results highlight broad networks of genes affecting fertility and calf survival through pathways involved in hormone signaling, energy balance, immunity, and cellular proliferation. The lethal haplotype and TRD analyses, in turn, underscore the importance of accounting for recessive lethals and semi-lethals that may remain undetected in traditional evaluations, yet substantially reduce conception success and increase pregnancy loss and early mortality. The integration of these approaches is particularly relevant in the context of genomic selection, where intense use of a limited number of superior sires can inadvertently increase the frequency of deleterious alleles if they are linked to favorable variants for production traits.

From a practical standpoint, the findings of this dissertation have direct implications for the design and management of genetic improvement programs in Nellore cattle. First, the genomic regions and candidate genes revealed by GWAS can be incorporated into genomic prediction pipelines, potentially improving the accuracy of breeding values for reproductive traits, which are typically measured with low precision. Markers located in or near these regions could be prioritized in custom genotyping arrays or used to refine selection indices that combine fertility, survival, and production traits. Second, the catalog of putative lethal haplotypes and TRD regions offers an immediate tool for managing matings: by avoiding or minimizing carrier-by-carrier matings for the most impactful haplotypes, breeding organizations can reduce the incidence of embryonic and fetal losses without compromising genetic gain. Implementing simple constraints or penalties on such matings in existing mating allocation software would represent a low-cost, high-impact intervention.

Moreover, the integration of TRD information into genomic evaluations and mating recommendations represents a promising frontier. Regions exhibiting strong TRD and associated with substantial underrepresentation of offspring genotypes can be used as early-warning indicators of underlying deleterious variation. As genotyping becomes increasingly routine, these signals can be monitored over time to prevent the

silent accumulation of harmful alleles in nucleus herds and commercial populations. In parallel, the knowledge generated by this study can contribute to more sustainable and ethically aligned breeding systems: reducing involuntary reproductive losses improves animal welfare and optimizes the use of natural resources in tropical beef production systems.

Despite these advances, several important research questions remain. The next step is to fine-map the most promising lethal haplotype and TRD regions using whole-genome sequence data. This would allow the identification of candidate causal variants, including SNPs, small insertions and deletions, and structural variants, and their functional annotation at base-pair resolution. Integrating transcriptomic data from reproductive tissues, embryos, and placentae would further clarify which genes and pathways are effectively dysregulated in carrier animals or affected conceptuses, moving from correlative genomic signals to mechanistic understanding. Another promising direction is the validation of key regions across multiple Nellore populations and other indicine breeds. Because many elite sires are used internationally, some lethal or semi-lethal alleles may be shared across breeds and countries, while others may be specific to particular selection histories or management systems. Cross-population analyses would help distinguish broadly relevant targets from population-specific variants and could inform international guidelines for the monitoring and control of deleterious alleles in *Bos taurus indicus* beef cattle.

Future studies would also benefit from integrating genomic information with more detailed phenotyping of reproductive function, including endocrine profiles, ultrasonographic measures of ovarian and uterine dynamics, embryo viability assessments, and environmental indicators such as temperature-humidity index and nutritional status. Linking genomic markers and haplotypes to intermediate phenotypes could help disentangle the pathways through which deleterious variants exert their effects and clarify gene–environment interactions that modulate reproductive outcomes in tropical conditions. Methodological developments are also needed to formally incorporate TRD and lethal haplotype information into routine genomic evaluations. This includes exploring penalization schemes for deleterious alleles in selection indices and simulating long-term consequences of different management strategies for deleterious variation under realistic breeding program scenarios. Such developments would support the transition from research-oriented TRD analyses to operational tools for genetic evaluation and mating design.