SÃO PAULO STATE UNIVERSITY
COLLEGE OF AGRICULTURAL AND VETERINARIAN SCIENCES
CAMPUS JABOTICABAL

RECOMBINATION EVENTS AND VARIANCE OF GAMETIC DIVERSITY ESTIMATION IN NELLORE CATTLE: IMPLICATIONS FOR BREEDING PROGRAMS

Patrícia Iana Schmidt
Master in Genetics and Animal Breeding
Animal Scientist

2025

SÃO PAULO STATE UNIVERSITY
COLLEGE OF AGRICULTURAL AND VETERINARIAN SCIENCES
CAMPUS JABOTICABAL

RECOMBINATION EVENTS AND VARIANCE OF GAMETIC DIVERSITY ESTIMATION IN NELLORE CATTLE: IMPLICATIONS FOR BREEDING PROGRAMS

Patrícia Iana Schmidt
Advisor: Profª. Dra. Lucia Galvão de Albuquerque
Co-advisor: Dr. Daniel Jordan de Abreu Santos
Co-advisor: Dr. Lúcio Flávio Macedo Mota

Thesis presented to the College of Agricultural and Veterinarian Sciences - São Paulo State University, Campus of Jaboticabal in partial fulfillment of requirements for the degree of Doctorate in Genetics and Animal Breeding.

2025

F464r

Fieno, Patrícia Iana Schmidt
Recombination events and variance of gametic diversity estimation in Nellore cattle: implications for breeding programs / Patrícia Iana Schmidt Fieno. -- Jaboticabal, 2025
156 p. : tabs.

Tese (doutorado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal
Orientadora: Lucia Galvão de Albuquerque
Coorientador: Daniel Jordan de Abreu Santos

1. Genômica. 2. Recombinação genética. 3. Diversidade genética. 4. Melhoramento genético. 5. Nelore. I. Título.

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

Registro de Impacto Esperado na Sociedade

Este estudo contribui para o avanço do melhoramento genético de bovinos da raça Nelore, gerando ferramentas que promovem maior eficiência, diversidade genética e sustentabilidade nos sistemas de produção de carne no Brasil.

Expected Societal Impact

This study supports the genetic improvement of Nellore cattle by providing tools that enhance efficiency, genetic diversity, and sustainability in Brazilian beef production systems.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: RECOMBINATION EVENTS AND VARIANCE OF GAMETIC DIVERSITY ESTIMATION IN NELLORE CATTLE: IMPLICATIONS FOR BREEDING PROGRAMS

AUTORA: PATRÍCIA IANA SCHMIDT
ORIENTADORA: LUCIA GALVÃO DE ALBUQUERQUE
COORIENTADOR: LÚCIO FLÁVIO MACÊDO MOTA
COORIENTADOR: DANIEL JORDAN DE ABREU SANTOS

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

Pesquisador Dr. DANIEL JORDAN DE ABREU SANTOS (Participaçao Virtual)
Departamento de Zootecnia / FCAV UNESP Jaboticabal

gov.br Documento assinado digitalmente
LENIRA EL FARO ZADRA
Data: 26/02/2025 10:51:44-0300
Verifique em https://validar.iti.gov.br

Pesquisadora Dra. LENIRA EL FARO ZADRA (Participaçao Virtual)
Instituto de Zootecnia / Sertãozinho/SP

gov.br Documento assinado digitalmente
GREGORIO MIGUEL FERREIRA DE CAMARGO
Data: 26/02/2025 12:46:11-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. GREGÓRIO MIGUEL FERREIRA DE CAMARGO (Participaçao Virtual)
Departamento de Zootecnia / Universidade Federal da Bahia (UFBA) - Salvador/BA

gov.br Documento assinado digitalmente
ANA FABRICIA BRAGA MAGALHAES
Data: 26/02/2025 09:23:09-0300
Verifique em https://validar.iti.gov.br

Profa. Dra. ANA FABRÍCIA BRAGA MAGALHÃES (Participaçao Virtual)
Departamento de Zootecnia / Universidade Federal dos Vales do Jequitinhonha e Mucuri - Diamantina/MG

Documento assinado digitalmente
gov.br
DANISIO PRADO MUNARI
Data: 26/02/2025 16:53:12-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. DANISIO PRADO MUNARI (Participaçao Virtual)
Departamento de Ciencias Exatas / FCAV UNESP Jaboticabal

Jaboticabal, 25 de fevereiro de 2025

ABOUT THE AUTHOR

PATRÍCIA IANA SCHMIDT – born on May 19, 1994, in Concordia, Santa Catarina, Brazil, daughter of Irdes Kochenborger Schmidt and Volnei Antonio Schmidt. She started her undergraduate in Animal Science in May 2013 at the Federal University of Pelotas, Campus of Capão do Leão. She was a scientific initiation scholarship holder from 2015 to 2018 under the supervision of Prof. Dra. Arione Augusti Boligon. During her undergraduate studies she also did internships at beef cattle farms. She obtained a bachelor's degree in animal science in February 2018. She started her Master of Science (MSc.) studies in the Graduate Program in Genetics and Animal Breeding at FCAV-UNESP, Campus of Jaboticabal, in March 2018, under the supervision of Prof. Dra. Lucia Galvão de Albuquerque and Dra. Ana Fabricia Braga Magalhães. Her M.Sc. studies have been granted with a scholarship from CAPES, from March 5, 2018, to February 28, 2019. She was a FAPESP scholarship holder (#2018/17812-1) from March 1, 2019, to July 31, 2020. She did an internship during November 1, 2019, to April 30, 2020, at the University of Maryland and at Animal Genomic Improvement Laboratory (AGIL/USDA) under the supervision of Dr. Li Ma, Dr. Paul VanRaden and Dr. John B. Cole, supported by FAPESP (#2019/10123-9). She started her Doctorate (Ph.D.) studies in the Graduate Program in Genetics and Animal Breeding at FCAV-UNESP, Campus of Jaboticabal on August 3, 2020, under the supervision of Prof. Dra. Lucia Galvão de Albuquerque, Dr. Daniel Jordan de Abreu Santos and Dr. Lúcio Flávio Macedo Mota. Her Ph.D. studies have been granted with a scholarship from CAPES, from August 3, 2020, to February 28, 2022. She was a FAPESP scholarship holder (#2021/09942-5) from March 1, 2022, to July 30, 2024.

"It is my most sincere wish that some of you will continue doing scientific work and maintain the ambition and determination to make a lasting contribution to science. But remember, you cannot hope to build a better world without improving individuals. To that end, each of us must work for our own improvement while also sharing a general responsibility for all of humanity."

Marie Curie

I dedicate this work entirely to God, who empowered and strengthened me.
To Him, to Whom I owe everything.
"There is no other way, my children. Either we know to find the Lord in our daily lives, or we will never find Him."
São Josemaria Escrivá

ACKNOWLEDGMENTS

Firstly, I thank God for everything He has done for me, for enlightening my path, and for helping me face difficult situations that have undoubtedly transformed me into a better person.

I am deeply grateful,

To my family, my husband Gustavo, whom I had the grace to meet during this challenging phase of my life, and to our son Rafael, the greatest gift of our lives. For strengthening me, for making me believe that it would be possible, and for helping me make this achievement real. For presence, support, love, patience and loyalty. For the willingness to build a family and share life with me.

To Father Ermínio, Sister Maria Augusta, and Lurdinha for their spiritual guidance, for welcoming me, and for guiding me with wisdom, patience, and love on the path of faith. For being instruments of God in my life and making me a better human being.

To my parents, Irdes e Volnei and my brother Lucas, for their unconditional support. Thank you for encouraging me to leave my comfort zone and go into the search for my professional realization. For motivating me every day, for believing in me, for never measuring efforts to help me, and for demonstrating every single day of my life the most pure and sincere feeling of love I have ever received.

To my special friend and co-worker Larissa S., who was always available to help me since my final internship, for professional encouragement, for the opportunities to work on other projects, for never measuring efforts to help me, for attention and especially, for loyal friendship.

To my friend Larissa B., thank you for all her help and understanding, for your trust, for the long work meetings, for the constructive conversations, for working with me on all DOGMA projects, and mainly for her loyal friendship.

To my great friends Natália C., Luciana, Paulo, Edinara, Pamela, and Heron, for their friendship and for making me feel better during this challenging time. For all the constructive conversations, the help, the advice, the special moments of happiness, and especially for the loyal and true friendship.

To my friends and colleagues, Prof. Dra. Ana Fabrícia, Prof. Dr. Ricardo, Dr. Lívia, Dr. Fabiana, and M.Sc. Diogo, for the knowledge shared, for believing in DOGMA's projects, for accepting to work with me, for including me in other projects and for the opportunities entrusted to me.

To my friends and colleagues Andrea, André, Júlio, Fernando, and Ana Maria, for the welcome in Araçatuba. Thank you for believing in my work, for trust, for constructive conversations, for the advice, and for the opportunity to work together at AgroPartners.

To my lab colleagues Leonardo, Gustavo, Daniel, Natália M., Sindy, João, Bruna, and Gabriela, who were always available to help me. I appreciate the opportunities to work on other projects and the support and attention given to the development of my projects.

To all the professors, staff and collaborators of FCAV/UNESP for all their teaching, availability and attention. Thank you for contributing to my professional growth and for helping me whenever I ask for help throughout all these years.

To my co-advisors, Daniel and Lucio, for giving me full support and attention during the development of this thesis and for putting all their knowledge within my reach. For their essential help from beginning to end.

To Prof. Dr. Li Ma from the University of Maryland, for all the support and attention given to me, even from a distance. For always responding to me quickly, for having provided his pipelines that made all the difference in the development of my projects.

To my advisor, Prof. Dra. Lucia, thank you for the opportunity to work in her research group since the beginning of my final internship. Thank you for your patience and understanding in both research and personal life. I particularly want to thank her for giving me the opportunity to join her laboratory and work on these exciting projects.

This study was financed in part by the São Paulo Research Foundation (Grants #2009/16118-5, #2017/10630-2, #2018/20026-8, #2019/10123-9, and #2021/09942-5).

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

SUMMARY

EVENTOS DE RECOMBINAÇÃO E ESTIMATIVA DA VARIÂNCIA DA DIVERSIDADE GAMÉTICA EM BOVINOS NELORE: IMPLICAÇÕES PARA PROGRAMAS DE MELHORAMENTO

RESUMO – A variância da diversidade gamética refere-se à variabilidade do potencial genético de um indivíduo que pode ser transmitido à sua prole por meio de seus gametas. Essa variância surge de dois processos genéticos fundamentais: amostragem mendeliana, que introduz aleatoriedade na segregação de alelos durante a meiose, e recombinação, que é um processo biológico fundamental que cria combinações de material genético. Os objetivos desta tese foram inicialmente calcular taxas de recombinação entre marcadores SNPs de painéis comerciais, identificar regiões de *hotspots* no genoma, construir um mapa de recombinação para machos Nelore, estimar parâmetros genéticos relacionados ao processo de recombinação, realizar estudo de associação genômica ampla (GWAS) para eventos de recombinação e identificar por meio de análises funcionais os genes candidatos próximos aos SNPs significativamente associados. Posteriormente, objetivou-se estimar a variância da diversidade gamética de todos os animais para diversas características economicamente importantes usando as taxas de recombinação entre marcadores SNPs previamente estimadas, bem como os efeitos de marcadores SNPs sobre tais características. Finalmente, a resposta à seleção para as características estudadas foi comparada considerando a seleção tradicional e a seleção considerando também a variância da diversidade gamética. Um total de 62.022 animais genotipados com painéis de baixa e média densidade foram imputados para um painel de alta densidade (409.029 marcadores SNP após controle de qualidade). As posições genômicas dos marcadores foram baseadas na montagem do genoma ARS-UCD1.2 *Bos taurus*. As características reprodutivas consideradas foram circunferência escrotal (SC), idade ao primeiro parto (IPP) e prenhez de novilhas (HP), reconcepção de novilhas (HR) e capacidade de permanência (STAY), e as características de crescimento foram peso corporal ao desmame (WW) e peso corporal ao ano (YW). Registros fenotípicos medidos em aproximadamente 1.100.000 animais Nelore de programas de melhoramento comercial foram usados. As taxas de recombinação foram medidas por meio de um método indireto usando pedigree e as informações de genotipagem. Foram extraídos pares progênie-pai do pedigree da população de gado Nelore estudada. Tanto o pai quanto a prole foram genotipados e faseados para inferir sobre eventos de recombinação para uma meiose paterna. As regiões de *hotspots* foram definidas como intervalos de SNPs com taxa de recombinação > 2,5 desvios padrão acima da média de recombinação. Os parâmetros genéticos para taxa de recombinação foram estimados usando dois modelos: 1) Usando o número médio de eventos de “crossovers” meióticos por reprodutor 2) Usando estes eventos como medidas

repetidas, uma vez que os reprodutores tiveram vários descendentes. Posteriormente, usando ambos os modelos, as análises de GWAS foram realizadas, considerando todos os “crossovers” localizados e somente as regiões de *hotspots,* para identificar genes candidatos e mecanismos biológicos que afetam a taxa de recombinação. Em análises funcionais, genes dentro de 100 kb a jusante e a montante de cada marcador SNP significativo foram investigados. Em análises subsequentes, aplicamos o mapa de recombinação construído para os machos, no software gamevar.f90 e estimamos a variância da diversidade gamética e o coeficiente de variação relativa para toda população, considerando características reprodutivas e de crescimento. Os genótipos foram faseados utilizando o software findhap.f90 v4 e para a obtenção dos efeitos dos marcadores para cada característica fenotípica estudada, um modelo assumindo homogeneidade de variância dos seus efeitos, ssGBLUP, foi usado. Como resultado, foram identificados 407.828 eventos de “crossovers” e 12.357 meioses paternas, que abrangem 33 Morgans para machos Nelore. Foram encontradas 2.039 regiões de *hotspots*, destacando-se o cromossomo 1 com o maior número destas regiões. As estimativas de herdabilidade foram de 0,24±0,06 e 0,17±0,09 para os modelos 1 e 2, respectivamente. Para o modelo 2 a repetibilidade estimada foi de 0,50±0,03. Por meio de análises GWAS, encontramos SNPs significativos para ambos os modelos (1 - 29SNPs; 2 - 192 SNPs) e genes semelhantes que podem impactar a recombinação de várias maneiras, seja participando diretamente do processo de recombinação homóloga ou influenciando indiretamente o ciclo celular, a estabilidade genômica e os processos de reparo do DNA. Nosso GWAS utilizando as regiões de hotspots da recombinação, identificou o gene *PRDM9* no cromossomo 1, em ambos os métodos, que está validado e diretamente relacionado com a recombinação meiótica. As estimativas de herdabilidade encontrados sugerem que os eventos de recombinação meiótica são um processo hereditário na população Nelore estudada. Por meio de nosso estudo, o primeiro mapa de recombinação para bovinos Nelore utilizando um painel comercial de SNP está sendo disponibilizado para a comunidade científica. A variância gamética foi identificada em todas as características, com BTA1 e BTA8 desempenhando papéis centrais na transmissão da variabilidade genética para características de crescimento e reprodução, respectivamente. O BTA1 teve o maior impacto na covariância gamética de YW, WWd e WWm, enquanto o BTA8 esteve associado a AFC, HP e HR, reforçando sua importância na regulação genética dessas características. A inclusão da variância gamética na seleção resultou em ganhos genéticos adicionais, com o maior aumento observado para HP (+8,41%) e STAY (+8,43%) sob REPD1.5. As características de crescimento apresentaram melhorias mais modestas, variando de 2,37% a 5,38%. O número de progênies necessário para capturar 90% da variância gamética variou entre 51 (SC e YW) e 125 (HP), destacando o potencial da seleção relativa em biotecnologias reprodutivas. Dado que este é o primeiro estudo na raça Nelore, nossos resultados revelam novos insights sobre a compreensão da recombinação e

aplicação da variância gamética nessa raça de alta importância econômica para o Brasil.

Palavras-chave: amostragem mendeliana, bovinos de corte, “crossover”, faseamento de genótipos, recombinação, variância gamética

RECOMBINATION EVENTS AND VARIANCE OF GAMETIC DIVERSITY ESTIMATION IN NELLORE CATTLE: IMPLICATIONS FOR BREEDING PROGRAMS

ABSTRACT – The variance of gametic diversity refers to the variability in the genetic potential of an individual that can be transmitted to its offspring through its gametes. This variance arises from two fundamental genetic processes: Mendelian sampling, which introduces randomness in allele segregation during meiosis, and recombination, an essential biological process that creates combinations of genetic material. The objectives of this thesis were to calculate recombination rates between SNP markers from commercial panels, identify hotspot regions in the genome, construct a recombination map for Nellore males, estimate genetic parameters related to the recombination process, conduct a genome-wide association study (GWAS) for recombination events, and identify candidate genes through functional analyses near significantly associated SNPs. Subsequently, the aim was to estimate the variance of gametic diversity for all animals across various economically important traits using the recombination rates between previously estimated SNP markers, as well as the effects of SNP markers on such traits. Finally, the response to selection for the traits studied was compared, considering traditional selection and selection based on the variance of gametic diversity. A total of 62,022 animals genotyped with low- and medium-density panels were imputed to a high-density panel (409,029 SNP markers after quality control). The genomic positions of the markers were based on the ARS-UCD1.2 Bos taurus genome assembly. The reproductive traits considered were scrotal circumference (SC), age at first calving (AFC), heifer pregnancy (HP), heifer rebreeding (HR), and stayability (STAY), while the growth traits were weaning weight (WW) and yearling weight (YW). Phenotypic records measured in approximately 1,100,000 Nellore animals from commercial breeding programs were used. Recombination rates were measured using an indirect method based on pedigree and genotyping information. Pairs of progeny-sire from the Nellore cattle pedigree were extracted, with both the sire and offspring genotyped and phased to infer recombination events for paternal meiosis. Hotspot regions were defined as SNP intervals with a recombination rate >2.5 standard deviations above the mean recombination rate. Genetic parameters for recombination rate were estimated using two models: 1) using the average number of meiotic crossover events per sire, and 2) using these events as repeated measures, as sires had multiple offspring. Subsequently, using both models, GWAS analyses were performed, considering all located crossovers and only the hotspot regions, to identify candidate genes and biological mechanisms affecting the recombination rate. In functional analyses, genes within 100 kb upstream and downstream of each significant SNP marker were investigated. In subsequent analyses, we applied the recombination map constructed for males in the software gamevar.f90 and estimated the variance of gametic diversity and the relative coefficient of variation for the entire population, considering reproductive and growth traits. Genotypes were phased using the software findhap.f90 v4, and to obtain the marker effects for each studied phenotypic trait, a model assuming variance homogeneity of their effects, ssGBLUP, was used. As a result, 407,828 crossover events and 12,357 paternal meiosis were identified, covering 33

Morgans for Nellore males. A total of 2,039 hotspot regions were found, with chromosome 1 presenting the highest number of such regions. The heritability estimates were 0.24±0.06 and 0.17±0.09 for models 1 and 2, respectively. For model 2, the repeatability estimate was 0.50±0.03. Through GWAS analyses, we found significant SNPs for both models (1 - 29 SNPs; 2 - 192 SNPs) and similar genes that may impact recombination in various ways, either by directly participating in the homologous recombination process or indirectly influencing the cell cycle, genomic stability, and DNA repair processes. Our GWAS using recombination hotspot regions identified the PRDM9 gene on chromosome 1, in both methods, which is validated and directly related to meiotic recombination. The heritability estimates values found that suggest that the meiotic recombination events are a heritable process in the Nellore population studied. Through our study of the first recombination map for Nellore cattle, an SNP panel is being made available to the scientific community. Gametic variance was identified across all traits, with BTA1 and BTA8 playing central roles in transmitting genetic variability for growth and reproductive traits, respectively. BTA1 had the greatest impact on the gametic covariance of YW, WWd, and WWm, whereas BTA8 was associated with AFC, HP, and HR, reinforcing their importance in genetic regulation. The inclusion of gametic variance in selection resulted in additional genetic gains, with the largest increase observed for HP (+8.41%) and STAY (+8.43%) under REPD1.5. Growth traits showed more modest improvements, ranging from 2.37% to 5.38%. The number of progenies required to capture 90% of the gametic variance varied between 51 (SC and YW) and 125 (HP), highlighting the potential of relative selection in reproductive biotechnologies. Since this is the first study in the Nellore breed, our results reveal new insights into the understanding of recombination and the application of gametic variance in this breed, which is so important for Brazil.

Keywords: beef cattle, crossover, gametic variance, genotype phasing, Mendelian sampling, recombination

CHAPTER 1 – General consideration

1.1 INTRODUCTION

Historically, livestock breeding efforts have focused on the selection of animals based on phenotypic values and, later, with advancements in statistical genetics, on estimated breeding values (EBVs). These EBVs reflect the additive genetic merit of an individual but do not account for variability introduced by the meiotic process, such as recombination and Mendelian sampling (VANRADEN, 2020). The breeding programs primarily aim to increase the frequency of favorable alleles, thereby improving the phenotypic mean of populations. However, EBVs represent only the average sum of additive effects of all genes, which ignores the variability caused by processes like Mendelian sampling, where alleles are randomly inherited by the progeny, leading to genetic diversity among full siblings (SEGELKE et al., 2014; TABET, 2022).

With the advent of large-scale genotyping technologies, the landscape of genetic evaluation has been transformed. Genomic selection (GS), introduced by Meuwissen et al. (2001) and later expanded by Schaeffer (2006), has allowed breeders to leverage dense molecular marker information to predict genomic estimated breeding values (GEBVs). This approach enables faster genetic progress and greater predictive accuracy compared to traditional EBV-based selection methods. GS allows breeders to utilize molecular marker data for more precise selection, even for individuals without phenotypic records, by estimating genome-wide marker effects (Meuwissen et al., 2001). However, despite the progress brought by GS, it still does not fully address the complexity of gametic diversity, particularly the influence of recombination and Mendelian sampling variance on the genetic potential of future progeny (COLE & VANRADEN, 2011).

Recombination, a key process in meiosis, plays an essential role in generating gametic diversity. It is the mechanism by which new genetic combinations are created, impacting the genetic variability of offspring. Recombination rates can be detected using various methods, including single-sperm genotyping (LU et al., 2012) and molecular assays (PAN et al., 2011), or indirectly through linkage disequilibrium maps and by computing the frequency of crossover events in genotyping data (MA et al., 2015; SHEN et al., 2018). Considering recombination data in genetic selection models enables breeders to

make more informed choices about the genetic potential of future offspring, which can significantly influence genetic gain (SANTOS et al., 2019). Maintaining genetic variability is especially important for Nellore cattle, given their crucial role in Brazil's extensive and varied production systems due to their adaptability to harsh environmental conditions.

To address this, recent advances in genomic selection have explored methods that incorporate gametic variance, which reflects the variability in genetic combinations that can arise from an individual's gametes. Santos et al. (2019) and (2020) have shown that considering gametic variance in breeding programs can enhance genetic progress by identifying individuals more likely to produce offspring with extreme breeding values. Methods such as Standard Deviation of Gamete Breeding Values (SDGBV) (SEGELKE et al., 2014) and Expected Maximum Haploid Reproductive Value (EMBV) (MÜLLER et al., 2018) have been developed to estimate the variability of gametic diversity, enabling more strategic selection decisions.

By incorporating gametic variance into selection models, breeders can select not only individuals with high genetic merit but also those that are likely to produce progeny with favorable genetic traits. Daetwyler et al. (2015) demonstrated that using the Optimal Haploid Value (OHV) method, which accounts for gametic combination, results in significantly higher genetic gains compared to traditional GS. This strategy can also help mitigate the negative effects of inbreeding, as it preserves greater genetic diversity within the population, ensuring its ability to adapt to future selection pressures (MALTECA et al., 2020; NIEHOFF et al., 2024a).

In the context of Nellore cattle, which are the main breed used in the tropical beef production system and the foundation of Brazil's beef export industry, maintaining genetic diversity through recombination and managing gametic variance is important for long-term sustainability. Despite the advances in other cattle breeds, such as Holstein and Jersey (MA et al., 2015; SHEN et al., 2018; SANTOS et al., 2019), there remains a notable research gap regarding the application of gametic variance in Nellore cattle breeding programs. As the largest producer and exporter of beef, Brazil stands to benefit significantly from integrating these methods into its breeding strategies for Nellore cattle, particularly given its diverse environmental challenges.

In this review, we will explore the current state of research on meiotic recombination, gametic diversity, Mendelian sampling variance, and their implications for breeding programs. We will also examine the practical challenges and prospects of incorporating these concepts into large-scale breeding programs, providing a comprehensive overview of their potential impact on the genetic improvement of this essential breed.

1.2 LITERATURE REVIEW

1.2.1 Meiotic recombination and Genetic Diversity

Recombination is a fundamental biological process that is essential for the creation of genetic diversity. During meiosis, homologous chromosomes exchange genetic material through crossover events, which shuffles alleles inherited from both parents. Thus, the meiosis process not only introduces novel genetic combinations in the offspring but also ensures proper chromosome segregation (BARTON & CHARLESWORTH, 1998; COOP & MYERS, 2007). In livestock, recombination plays a pivotal role in enhancing genetic diversity, which is essential for the success of selective breeding programs aimed at improving economically important traits.

The recombination events can be assessed by comparing the haplotypes of an individual in relation to those of its parents (TAO et al., 2018) using direct and indirect measurements. Direct methods measure the recombination events by examining molecular evidence related to specific stages of meiosis, such as Single-Sperm Genome-Wide Genotyping (LU et al., 2012; WANG et al., 2012) or Molecular Assay that tracks markers of recombination activity (PAN et al., 2011). While indirect methods rely on patterns of Linkage Disequilibrium (LD) to build recombination maps by identifying local patterns and computing the frequency of crossover events between SNPs in phased genotyping data (MOURESAN et al., 2019). However, the indirect method relies on populations being genetically homogeneous or having conserved recombination rates across generations, as observed in closely related breeds or structured breeding programs. When these assumptions are not met, pedigree data becomes essential to ensure accurate inference of recombination rates (MA et al., 2015; SHEN et al., 2018). Thus, pedigree-based analyses offer a reliable method for tracing recombination events through family lines, allowing accurate estimates of recombination events in genetically complex populations.

Indirect methods have the advantage of estimating recombination rates at the population level compared with direct methods that rely on a group of individuals (SHEN et al., 2018), making them more efficient for estimating gametic variance. Furthermore, Zhou et al., 2018 evaluated the reliability of pedigree-based and single sperm typing approaches, compared recombination patterns between sperm and live-born offspring, and observed similar

recombination patterns across the genome, validating the quality of pedigree-based results. Despite several studies that have shown the importance of understanding the pattern and inheritance of recombination rates and how to explore selection methods that can release favorable alleles, contributing to increasing genetic gain, no specific studies were found with the Nellore breed or even for any Zebu breeds.

In Nellore cattle, recombination studies are especially relevant due to the breed's extensive use in beef production in Brazil. Representing a large proportion of the cattle population in tropical regions, this breed offers an important opportunity to understand how recombination influences genetic diversity and supports the refinement of breeding strategies. Recombination plays a key role in generating gametic diversity, a fundamental driver of genetic improvement in breeding populations (WENG et al., 2019). Variation in recombination rates has been documented across species, breeds, and even between sexes, with both genetic and environmental factors exerting considerable influence on the process (OLLIVIERZ, 1995; BROMAN et al., 1998; SARDELL & KIRKPATRICK, 2020; PEÑALBA et al., 2020; FURMAN et al., 2020; PETERSON & PAYSEUR, 2021).

Understanding the genetic basis of recombination can provide breeders with the tools to maximize genetic gains through optimized mating strategies that promote favorable allele combinations (KADRI et al., 2022; BREKKE et al., 2022). The advancements in genomic technologies, including genome-wide association studies (GWAS), have enhanced our understanding of genetic architecture controlling recombination mechanisms. This knowledge can be applied to improve the accuracy and efficiency of breeding programs in Nellore and other cattle breeds.

Recombination in eukaryotes occurs during the first phase of meiosis, where homologous chromosomes align and experience crossover events. This process is tightly regulated and ensures that at least one crossover in one of the sister chromatids per chromosome pair occurs to maintain genetic diversity and chromosome stability (KAUPPI et al., 2004). The molecular processes involved in recombination include several key proteins, such as those encoded by the *RNF212* and *MSH4* genes, which have been shown to regulate crossover formation and the repair of double-strand breaks (BREKKE et al., 2022). In

addition, several loci, such as *RNF212, PRDM7, PRDM9, NEK9, MLH3, MSH4, MSH5, REC114, REC8, FMN1, CPLX1* and *SYCP2*, have been identified as key regulators of recombination frequency in mammals (MA et al., 2015; KADRI et al., 2016; QANBARI & WITTENBURG, 2020; BREKKE et al., 2022) and for example *MSH4*, *SMC3* and *CEP55*, which affect recombination rate in females only (MA et al., 2015).

In cattle, as in many other mammals, recombination events are not randomly distributed across the genome. Instead, recombination events are concentrated in specific genomic regions known as hotspots, where crossovers are more likely to occur. These hotspot regions are essential for shaping the genomic landscape and influencing patterns of inheritance in cattle populations (KAUPPI et al., 2004; PAIGEN & PETKOV, 2010). In humans, recombination hotspots are typically clustered in regions that are mainly controlled by the *PRDM9* gene, which determines where recombination events are likely to occur (PAIGEN & PETKOV, 2010). Understanding the location and activity of these hotspots provides valuable insights into how recombination contributes to genetic diversity.

Studies have characterized and explored the role of the *PRDM9* gene in meiotic recombination. This gene is essential for regulating recombination hotspots during meiosis in many mammals, including cattle. By determining the locations of hotspots, *PRDM9* guides where DNA breaks occur, facilitating the exchange of genetic material between homologous chromosomes (PAIGEN & PETKOV, 2018). In cattle, the diversity of *PRDM9* sequences, particularly in its zinc finger region, affects the pattern and distribution of recombination hotspots (ZHOU et al., 2018). Multiple *PRDM9* variants have been identified, showing differences in the number of zinc finger repeats across breeds, ranging from 6 to 9 repeats (AHLAWAT et al., 2017; ZHOU et al., 2018). These variations contribute to the diversification of recombination hotspots, which enhances genetic variability, since it is an important factor in domestication and evolutionary processes (ZHOU et al., 2018; SEROUSSI et al., 2019).

Beyond its role in determining hotspots, *PRDM9* acts as an epigenetic regulator during meiosis. Through its PR/SET domain, it modifies chromatin by adding chemical marks, such as the trimethylation of histone H3 at lysine 4 (H3K4me3) and 36 (H3K36me3). These marks alter the chromatin structure,

making it more accessible for the formation of double-strand breaks (DSBs) required for recombination (PAIGEN & PETKOV, 2018; GREY et al., 2018). *PRDM9* also interacts with proteins like *EHMT2* and *CDYL*, anchoring hotspots to the chromosomal axis and ensuring accurate formation and repair of DSBs. This dual role epigenetic modification and structural anchoring allows *PRDM9* to maintain genetic diversity and guide evolutionary changes in cattle populations (DI TULLIO et al., 2022; ZHOU et al., 2018).

PRDM9's influence extends further, impacting fertility and speciation in mammals. In hybrids of different subspecies of cattle, such as Bos taurus × Bos indicus crosses, incompatibility between *PRDM9* alleles has been linked to reduced recombination rates and male infertility (SEROUSSI et al., 2019). Similar effects have been observed in other species; for instance, PRDM9 is the only known hybrid sterility gene in mice (AHLAWAT et al., 2017; PAIGEN & PETKOV, 2018). In cattle, specific *PRDM9* variants have been shown to drive differences in hotspot usage, which directly affects recombination rates, and the genetic diversity passed to offspring (SANDOR et al., 2012; ZHOU et al., 2018). These findings underscore *PRDM9's* central role in genetic diversity and evolutionary processes across species.

Studies have demonstrated that recombination rates vary widely between sexes, with males and females often exhibiting distinct patterns of crossover distribution. For instance, in Holstein cattle, males tend to have higher recombination rates near the telomeres, while females exhibit a more even distribution along the chromosomes (MA et al., 2015; SHEN et al., 2018). This trend was also observed in other cattle breeds, such as Norwegian Red, Jersey, Ayrshire, and Brown Swiss, where sex-specific differences in recombination rates were reported (SHEN et al., 2018; BREKKE et al., 2023a). The sex-specific differences are also seen in other species, such as pigs and Atlantic salmon, where female recombination rates were consistently higher than males across most chromosomes (BREKKE et al., 2022; BREKKE et al., 2023b).

Studies have explored the difference in recombination rates along chromosomes. In many species, recombination is suppressed near centromeres and elevated near telomeres (NACHMAN, 2002; KAUPPI et al., 2004). However, the patterns of recombination are highly variable. In humans, mice and Atlantic salmon for example, females generally exhibit higher recombination rates than

males (NACHMAN, 2002; ZHOU et al., 2018; BREKKE et al., 2023b). In cattle, the opposite trend is observed, with males displaying higher overall recombination rates than females, a pattern especially pronounced in the sub telomeric regions of autosomes (MA et al., 2015; ZHOU et al., 2018; QANBARI & WITTENBURG, 2020; BREKKE et al., 2023a). Studies involving large-scale pedigree analyses and sperm genomics have revealed that male recombination rates are about 10% higher than female rates in cattle (MA et al., 2015; SANDOR et al., 2012). This sex-specific difference, particularly in cattle, contrasts with the patterns observed in humans and other mammals, where females typically show higher recombination rates, reflecting an evolutionary divergence in the regulation of recombination across species (ZHOU et al., 2018).

Several genetic factors contribute to these variations, with the *PRDM9* gene being a key regulator of recombination hotspots in many species (PAIGEN & PETKOV, 2018; ZHOU et al., 2018). In cattle, studies have shown that different alleles of *PRDM9* lead to variation in hotspot locations, particularly at the chromosome ends (sub telomeric regions), further influencing the overall recombination landscape (SANDOR et al., 2012; MA et al., 2015; ZHOU et al., 2018). This gene plays a key role in coordinating the formation of double strand breaks that initiate recombination and has been linked to the gradual loss of recombination hotspots over generations (ZHOU et al., 2018). Additionally, quantitative trait loci (QTL) such as *REC8* and *RNF212* also significantly influence recombination rates, particularly in males (SANDOR et al., 2012; WENG et al., 2014). These findings highlight the complex interplay between genetic and environmental factors in shaping recombination patterns across species, sexes and genomic regions, which contributes to genetic diversity and evolution and has important implications for breeding and selection strategies in livestock (SANDOR et al., 2012).

To investigate the genetic architecture of recombination rates on the X chromosome in cattle, Zhang et al. (2020) highlighting significant differences between males and females. In males, recombination is restricted to the pseudo-autosomal region (PAR), while in females, recombination occurs across the entire X chromosome. Using a large dataset of over 100,000 genotyped animals, this study found that recombination rates in the PAR were six times higher in males compared to females. Additionally, the heritability of recombination rates on the

X chromosome was estimated to be 0.024 in females and 0.011 in males, indicating stronger genetic control of this trait in females. The authors also observed a significant correlation between global recombination rates and X chromosome recombination rates in females, suggesting shared genetic control. These findings provide valuable insights for breeding programs aiming to enhance genetic diversity by optimizing recombination rates, particularly in female populations.

The study by Brekke et al. (2023b) is one of the first to investigate the genetic variation in individual meiotic crossover rates and their distribution specifically in Atlantic salmon. The research explored the extreme sex differences in recombination landscapes, where males predominantly exhibited crossovers in telomeric regions, whereas females showed a more even distribution across chromosomes, particularly near centromeres. Using a dataset from more than 5,000 full-sib families genotyped on ~35,000 SNP markers, this study found that female recombination led to approximately 1.6 times more crossovers than males, and females also exhibited eight times more genetic shuffling due to the broader distribution of crossovers. Despite these findings, no significant quantitative trait loci (QTLs) for either crossover count, or shuffling were identified, indicating a potentially polygenic nature for recombination in Atlantic salmon (BREKKE et al., 2023b).

The heritability of recombination rates in cattle has also been a subject of significant research, with studies consistently showing that recombination is a heritable trait. Heritability of recombination rates refers to the proportion of variation in crossover frequency that can be attributed to genetic factors. In cattle, studies show estimates ranging from 0.10 to 0.30 depending on the population and the method used to estimate it (KONG et al., 2004; SANDOR et al., 2012; WENG et al., 2014; WENG et al., 2019; KADRI et al., 2022). The heritability of recombination rates in pigs, as estimated by Brekke et al. (2022), was found to be slightly lower than that in cattle, ranging from 0.05 to 0.11. Similar values were reported in Atlantic salmon, ranging from 0.06 to 0.11 by Brekke et al (2023b). These findings suggest that recombination rates can be modified through selective breeding, providing a potential avenue for breeders to enhance genetic diversity in their herds.

Repeatability is another key concept in understanding recombination rates. Repeatability refers to the consistency of the expression of a trait in different periods of the life of the same individual. Kadri et al. (2022) found repeatability estimates for recombination rates in cattle ranging between 0.30 and 0.40, indicating that recombination frequency remains relatively stable across generations. This consistency is valuable for breeders, as it allows for obtaining more accurate predictions of recombination outcomes in progeny, improving the efficiency of breeding programs. In breeds as Nellore, where the long-term sustainability of breeding programs is critical, the stability of recombination rates could help optimize mating strategies.

Another way to explore recombination is through GWAS, which has played an important role in identifying the genetic loci associated with recombination rates in cattle. These studies have provided valuable insights into the genetic architecture of recombination, revealing several key genes that regulate crossover formation. One of the most significant findings from GWAS in cattle is the identification of the *RNF212* gene, which plays a critical role in regulating crossover events during meiosis (BREKKE et al., 2022, KADRI et al., 2022). The activity of *RNF212* is essential for ensuring the proper distribution of recombination events across the genome, and variations in this gene have been linked to differences in recombination rates between individuals.

In a study on New Zealand dairy cattle, Kadri et al. (2022) identified significant associations on the X-chromosome, with the *PPP4R3C* gene showing a strong influence on recombination rates in males. This gene is involved in the repair of double strand breaks during meiosis, further emphasizing its importance in regulating recombination. The study also found that the contribution of the X-chromosome to the heritability of recombination was higher in males (approximately 19.9%) compared to females, where the effect was negligible (KADRI et al., 2022). Similarly, Brekke et al. (2022) identified the *PRDM7* and *MEI1* genes as novel loci associated with recombination rates in pigs. These findings suggest that genetic control of recombination is conserved across species, with certain genes, such as RNF212 and *SYCP2*, playing a universal role in regulating crossover events in both cattle and pigs. The identification of these loci has important implications for breeding programs, as they offer

potential targets for gene editing or genomic selection to improve genetic diversity.

Understanding the genetic basis of recombination and its heritability has significant implications for cattle breeding programs. By incorporating recombination data into genomic evaluations, breeders can enhance the accuracy of their selection decisions and increase the genetic diversity of their herds (JENKO et al., 2015). This is particularly important for traits that are influenced by multiple genes, as recombination can break up unfavorable linkage phases and promote the spread of beneficial alleles (BATTAGIN et al., 2016).

The ability to predict recombination outcomes in progeny also allows breeders to design more effective mating strategies. For example, by selecting animals with higher recombination rates, breeders can increase the likelihood of producing offspring with novel genetic combinations, which can enhance the genetic progress of economically important traits (KADRI et al., 2022).

Looking ahead, the integration of recombination data into genomic selection models represents an exciting opportunity for the livestock industry. Advances in genomic technologies, such as whole-genome sequencing and gene editing, offer the potential to manipulate recombination rates directly, allowing breeders to maximize genetic gains while maintaining genetic diversity (BREKKE et al., 2022). However, challenges remain in incorporating these data into routine breeding programs, particularly given the sex-specific differences in recombination rates and the complex nature of the trait. Still, leveraging recombination data, cattle breeding programs can increase the efficiency of selection, promote genetic diversity, and ensure the long-term sustainability of economically important traits.

1.2.2 Gametic diversity and applications in breeding programs

The concept of gametic diversity variance has been explored, particularly in dairy cattle species, where it plays a key role in improving the prediction of offspring performance, genetic gain and ensuring long-term genetic diversity. This variance arises from two fundamental genetic processes: Mendelian sampling, which introduces randomness in allele segregation during meiosis, and recombination, which creates novel combinations of genetic material. These

processes ensure that each gamete produced by an individual carries a unique set of alleles, leading to variation among offspring, even in full-sib families (KAUPPI et al., 2004; BARTON & CHARLESWORTH, 1998).

Genomic selection has become a standard tool for predicting an individual's genetic potential based on Genomic Estimated Breeding Values (GEBVs). However, GEBVs do not account for the variance that arises from gametic diversity, limiting their ability to predict the range of possible genetic outcomes in the progeny (COLE & VANRADEN, 2011). This limitation has prompted researchers to incorporate gametic diversity variance into selection models, aiming to optimize genetic gains while maintaining genetic diversity across generations (VANRADEN, 2020).

The preservation of genetic variability is fundamental for the long-term sustainability of breeding programs, particularly in populations subjected to intense selection pressure. The variance of gametic diversity plays an essential role in maintaining this variability by ensuring that a wide range of genetic combinations is available for selection in future generations. Studies such as Santos et al. (2020) have demonstrated that selection based solely on GEBVs can lead to a reduction in genetic diversity over time, which can, in turn, increase the risk of inbreeding depression and reduce the adaptability of the population.

VanRaden (2020) and Lozada-Soto et al. (2021) further emphasize the importance of balancing genetic gain with the maintenance of genetic diversity in genomic selection programs. In commercial cattle populations, where high selection intensity is often applied to maximize production traits, incorporating gametic diversity variance allows breeders to retain genetic variability, which is essential for the adaptability and long-term health of the herd (LOZADA-SOTO et al., 2021; VANRADEN, 2020).

Some studies have explored the concept of gametic variance and its implications for genetic selection. Below, we explore the key studies, highlighting their methodologies, key findings, and differences in approach. This comparison provides a clear overview of how gametic diversity has been applied in various studies, showing the benefits and potential advances in simulation-based and population-specific research.

A notable study by Segelke et al. (2014), conducted on German Holstein cattle, explored the use of standard deviation of gamete breeding values

(SDGBV) to estimate the potential variability in offspring. Working with a dataset of 74,353 Holstein cattle, the researchers used phased genotypes and recombination data to calculate the SDGBV. The results demonstrated substantial differences in SDGBV between individuals, highlighting that even bulls with high GEBVs could produce offspring with widely varying genetic values. Bulls with higher SDGBV were found to be more likely to produce offspring with extreme breeding values, making them ideal candidates for long-term genetic improvement programs. This study reinforced the importance of selecting individuals not just based on GEBVs but also considering their potential to contribute greater genetic diversity to the population.

An alternative approach to utilizing the haplotype diversity was introduced by Daetwyler et al. (2015), in a simulation study focusing on dairy cattle. Instead of solely relying on GEBVs, this study proposed the Optimal Haploid Value (OHV), which seeks to select the best possible haplotypes that an individual could pass on to his offspring. The methodology involved simulating multiple generations of breeding and comparing OHV to traditional GEBV-based models. The results were striking, showing that the OHV method resulted in up to 0.6 standard deviations more genetic gain than GEBV-based selection while also helping to preserve genetic diversity over time. This simulation-based study emphasized the potential of methods that did not incorporate gametic diversity variance properly. However, they have shown how estimating haplotype combinations could be a useful tool for breeding programs where maximizing genetic gain while maintaining diversity is a priority.

In another significant study, Bonk et al. (2016), also working with German Holstein cattle, focused on estimating the within-family genetic variation caused by Mendelian sampling. Using detailed genotypic and recombination data, the researchers accurately calculated the variance in breeding values within families. The study found that incorporating Mendelian sampling variance into selection models improved the prediction of offspring performance, particularly for traits with moderate to high heritability. This approach provided a more nuanced understanding of how selection based solely on GEBVs might overlook important within-family variability, which could be critical for enhancing long-term genetic gain and maintaining genetic diversity.

Using a maize population in a simulation study, Goiffon et al. (2017) introduced the Optimal Population Value (OPV) selection, a novel genomic selection method designed to improve long-term responses in breeding programs. OPV differs from traditional selection methods by focusing not just on the individual breeding values but on the collective value of a group of individuals, maximizing the genetic potential of the population. Specifically, OPV identifies the set of individuals that, when selected together, optimize the diversity and complementarity of their haplotypes, ensuring a more favorable combination of alleles is passed on to future generations. This contrasts with approaches that focus solely on the best-performing individuals based on their genomic breeding values (GEBVs). The study compared OPV with standard Genomic Selection (GS), Weighted Genomic Selection (WGS), Optimal Haploid Value (OHV), and Genotype Building (GB). Through simulations, OPV proved more effective in maintaining genetic variability and achieving higher long-term genetic gains, outperforming other methods in 65.7% of the simulations. This makes OPV a powerful strategy for improving breeding programs by selecting complementary groups of individuals rather than just the highest-ranked ones (GOIFFON et al 2017).

Müller et al. (2018) further advanced the field with the introduction of the Expected Maximum Haploid Breeding Value (EMBV), an alternative selection criterion designed to evaluate the potential genetic contribution of an individual's best haplotypes. The study, which used simulated cattle populations, compared the performance of EMBV over multiple selection cycles. After five cycles, EMBV produced higher long-term genetic gains and preserved more genetic diversity than GEBV alone. This work demonstrated that EMBV could be a valuable tool for recurrent genomic selection, particularly in breeding programs where both genetic gain and the maintenance of genetic variability are priorities.

The concept of gametic diversity variance was developed to quantify the variability in genetic potential transmitted by gametes from an individual to its offspring. Santos et al. (2019) advanced this concept by proposing a formal statistical approach to estimate this variance, which is directly related to Mendelian sampling variance. The variance is calculated by considering the recombination and transmission of alleles at all quantitative trait loci (QTL), accounting for the probability of allele transmission and the genetic distance

between loci. By doing so, this method captures the genetic variability within an individual's gametes, which traditional breeding values fail to address. Santos et al. demonstrated that gametic diversity variance can be reliably estimated using single nucleotide polymorphism (SNP) marker data from routine genomic evaluations, with accuracy levels ranging from 0.49 to 0.85, depending on the model and scenario based on trait heritability and its number of QTLs. This approach allows the identification of individuals more likely to produce progeny with extreme breeding values, optimizing the selection process (SANTOS et al, 2019).

One of the major contributions of Santos et al. (2019) was demonstrating the potential of using gametic diversity variance in selection indices, particularly through the introduction of relative selection indices, named by them as the relative predicted transmitting ability (RPTA). These indices are relative to selection to be applied in the future generation, combining the expected value and a component derivate from the individual gametic variability. By incorporating the variance of gametic diversity into selection programs, the study showed that genetic progress could be improved by up to 16% over 10 generations when compared to traditional selection methods. The study also highlighted the ability of marker panels, even those with moderate densities, to predict gametic variance with sufficient accuracy, making the approach practical for large-scale genomic selection. This represents a significant advancement in methods that only focus on breeding values without considering the variability among gametes. The application of gametic variance to individual selection indices, such as RPTA, enhances the ability to select animals with a higher likelihood of producing genetically diverse and superior offspring, thereby maximizing long-term genetic gains (SANTOS et al., 2019; SANTOS et al., 2020).

Variance is a statistical parameter related to probability distributions. The concept of gametic variance has been explored in only a few studies, most notably by Segelke et al. (2014), Bonk et al. (2016), and Santos et al. (2019). In the latter, the parameter of gametic diversity variance is statistically grounded—through a formal statistical formula—rooted in quantitative genetics and first applied to individual selection indices. Santos et al. (2019) introduced the concept of "Gametic Diversity Variance," outlining the theoretical and statistical principles

behind gametic haplotype combinations with a focus on using statistical parameters to optimize selection based on gametic diversity.

Other studies, such as Daetwyler et al. (2015) and Müller et al. (2018), simulated genetic diversity and identified optimal haplotype combinations but neglected variance-related statistical parameters. As a result, selection based on these criteria shows lower genetic gain than methods that incorporate gametic diversity variance as a key statistical parameter.

Further advancing the understanding of gametic diversity variance, Bijma et al. (2020) conducted a study on Holstein-Friesian cattle, directly modeling GEBVs and gametic GEBV standard deviations. This study, varying only the weight of the gametic standard deviation, developed ten different selection indices that considered both GEBVs and gametic variance. The results showed that, including gametic variance in the selection process, genetic gains increased by up to 36% in top-ranking individuals while also maintaining a greater level of genetic diversity compared to models that only used GEBVs. This study demonstrated that accounting for the variability of gametes not only enhances immediate genetic progress but also contributes to preserving genetic diversity, which is important for the sustainability of breeding programs over multiple generations.

Tabet's (2022) investigated the use of Mendelian Sampling (MS) and gametic variance ($\sigma^2_{gametic}$) in selection programs. Through simulations using AlphaSimR, Tabet (2022) evaluated the impact of these selection schemes on genetic gain, genetic variability, and inbreeding rates over ten generations. The results showed that while the True Breeding Value (TBV) program achieved the highest cumulative genetic gain, the MS and $\sigma^2_{gametic}$ programs led to higher rates of long-term genetic gain, with the MS program showing 8.5% more gain in later generations compared to the TBV program. Additionally, the combination of TBV and $\sigma^2_{gametic}$ maintained higher additive genetic variance, approximately 3% greater than the TBV-only program, while also reducing inbreeding rates. These findings indicate that incorporating gametic variance into selection strategies can promote genetic variability and sustain genetic improvement over multiple generations.

Recently, two studies developed by Niehoff et al. (2024a) and Niehoff et al. (2024b), with distinct focuses and approaches, investigated improving breeding strategies through Mendelian sampling variances (MSV). The study by Niehoff et al. (2024a) introduces a new selection criterion called ExpBVSelGrOff (expected breeding value of selected grandprogeny), which incorporates information from two future generations into breeding decisions. This method optimizes mate selection by considering not only the breeding values (BV) of the individuals but also the potential genetic variance in their grand offspring. Through stochastic simulations, the study demonstrated that this approach leads to a 5.8% higher genetic gain and retains 25% more genetic variance over five generations compared to traditional selection methods. This finding underscores the importance of factoring in long-term genetic variability in breeding programs to ensure sustained genetic progress and maintain genetic diversity without increasing inbreeding.

In contrast, Niehoff et al. (2024b) focuses on an analytical approach for predicting the additive genetic variance of descendants based on the Mendelian sampling variances (MSV) of their ancestors. This study presents a novel method for estimating genetic variance in future generations, validated through simulations with both maize and cattle populations. The predictive correlations between the model and actual genetic variance were remarkably high, ranging from 0.90 to 0.97. This approach emphasizes the value of accurately predicting genetic variance to aid breeders in long-term planning, ensuring that future breeding populations retain both genetic diversity and potential for genetic improvement. In summary, while Niehoff et al. (2024a) introduces a practical selection tool for optimizing mate selection by looking two generations ahead, Niehoff et al. (2024b) provides an analytical framework for predicting genetic variance in future generations, both offering complementary insights into long-term genetic progress and sustainability.

The computational challenge of estimating gametic diversity variance for large populations has also been addressed through tools like Gamevar.f90, developed by Santos et al. (2020). Although this tool was not applied specifically to Nellore cattle, it has proven effective in other livestock populations, particularly in dairy breeds. Gamevar.f90 allows breeders to estimate the variance in gametic breeding values by analyzing phased genotypes and recombination data. This

software provides a practical solution for breeders looking to incorporate gametic diversity into their selection strategies, identifying individuals capable of producing offspring with extreme genetic values, thereby optimizing both genetic gain and genetic diversity.

Despite extensive research on gametic diversity variance in dairy cattle, there are currently no studies that specifically focus on Nellore cattle. This gap in the literature highlights the relevance and potential impact of applying gametic diversity variance in Nellore breeding programs, particularly given the breed's in tropical beef production. As Nellore cattle face unique environmental challenges, such as heat stress and parasite resistance, maintaining genetic diversity is critical for ensuring their long-term productivity and adaptability. The application of gametic diversity variance in Nellore cattle could provide a valuable tool for improving the genetic potential to produce this breed while preserving its genetic diversity, which could be a key factor in sustaining Brazil's position as the world's leading beef exporter.

1.2.3 References

ABIEC. Brazilian Livestock Profile Contents. (2023).

AHLAWAT, S., DE, S., SHARMA, P., SHARMA, R., ARORA, R., KATARIA, R. S., & SINGH, R. K. Evolutionary dynamics of meiotic recombination hotspots regulator PRDM9 in bovids. **Molecular genetics and genomics**, v. 292, p. 117-131, 2017.

ALBUQUERQUE, L. G., JÚNIOR, FERNANDES JÚNIOR, G.A., CARVALHEIRO, R. Beef Cattle Genomic Selection in Tropical Environments. In: **Proc. Assoc. Advmt. Anim. Breed. Genet**. 2017. p. 255-263.

BARTON, N. H.; CHARLESWORTH, B. Why sex and recombination? **Science**, v. 281, n. 5385, p. 1986-1990, 1998.

BATTAGIN, M., GORJANC, G., FAUX, A. M., JOHNSTON, S. E., & HICKEY, J. M. Effect of manipulating recombination rates on response to selection in livestock breeding programs. **Genetics Selection Evolution**, v. 48, p. 1-12, 2016.

BIJMA, P., WIENTJES, Y. C., & CALUS, M. P. Breeding top genotypes and accelerating response to recurrent selection by selecting parents with greater Gametic variance. **Genetics**, v. 214, n. 1, p. 91-107, 2020.

BONK, S., REICHELT, M., TEUSCHER, F., SEGELKE, D., & REINSCH, N. Mendelian sampling covariability of marker effects and genetic values. **Genetics Selection Evolution**, v. 48, n. 1, p. 36, 2016.

BREKKE, C., BERG, P., GJUVSLAND, A. B., & JOHNSTON, S. E. Recombination rates in pigs differ between breeds, sexes and individuals, and

are associated with the RNF212, SYCP2, PRDM7, MEI1 and MSH4 loci. **Genetics Selection Evolution**, v. 54, n. 1, p. 33, 2022.

BREKKE, C., JOHNSTON, S. E., GJUVSLAND, A. B., & BERG, P. Variation and genetic control of individual recombination rates in Norwegian Red dairy cattle. **Journal of Dairy Science**, v. 106, n. 2, p. 1130-1141, 2023a.

BREKKE, C., JOHNSTON, S. E., KNUTSEN, T. M., & BERG, P. Genetic architecture of individual meiotic crossover rate and distribution in Atlantic Salmon. **Scientific Reports**, v. 13, n. 1, p. 20481, 2023b.

BROMAN, K. W., MURRAY, J. C., SHEFFIELD, V. C., WHITE, R. L., & WEBER, J. L. Comprehensive human genetic maps: individual and sex-specific variation in recombination. **The American Journal of Human Genetics**, v. 63, n. 3, p. 861-869, 1998.

CARVALHEIRO, R. Genomic selection in Nellore cattle in Brazil. In: **Proceedings of the 10th World Congress on Genetics Applied to Livestock Production**. Vancouver, BC: Volume Species Breeding Beef Cattle, 2014. p. 17-22.

COLE, J. B.; VANRADEN, P. M. Use of haplotypes to estimate Mendelian sampling effects and selection limits. **Journal of Animal Breeding and Genetics**, v. 128, n. 6, p. 446-455, 2011.

COOP, G., & MYERS, S. R. Live hot, die young: transmission distortion in recombination hotspots. **PLoS genetics**, v. 3, n. 3, p. e35, 2007.

DAETWYLER, H. D., HAYDEN, M. J., SPANGENBERG, G. C., & HAYES, B. J. Selection on optimal haploid value increases genetic gain and preserves more genetic diversity relative to genomic selection. **Genetics**, v. 200, n. 4, p. 1341-1348, 2015.

DI TULLIO, F., SCHWARZ, M., ZORGATI, H., MZOUGHI, S., & GUCCIONE, E. The duality of PRDM proteins: epigenetic and structural perspectives. **The FEBS journal**, v. 289, n. 5, p. 1256-1275, 2022.

FURMAN, B. L., METZGER, D. C., DAROLTI, I., WRIGHT, A. E., SANDKAM, B. A., ALMEIDA, P., & MANK, J. E. Sex chromosome evolution: so many exceptions to the rules. **Genome biology and evolution**, v. 12, n. 6, p. 750-763, 2020.

GOIFFON, M., KUSMEC, A., WANG, L., HU, G., & SCHNABLE, P. S. Improving response in genomic selection with a population-based selection strategy: optimal population value selection. **Genetics**, v. 206, n. 3, p. 1675-1682, 2017.

GREY, C., BAUDAT, F., & DE MASSY, B. *PRDM9*, a driver of the genetic map. **PLoS genetics**, v. 14, n. 8, p. e1007479, 2018.

JENKO, J., GORJANC, G., CLEVELAND, M. A., VARSHNEY, R. K., WHITELAW, C. B. A., WOOLLIAMS, J. A., & HICKEY, J. M. Potential of promotion of alleles by genome editing to improve quantitative traits in livestock breeding programs. **Genetics Selection Evolution**, v. 47, p. 1-14, 2015.

KADRI, N. K., HARLAND, C., FAUX, P., CAMBISANO, N., KARIM, L., COPPIETERS, W., & DRUET, T. Coding and noncoding variants in HFM1, MLH3, MSH4, MSH5, RNF212, and RNF212B affect recombination rate in cattle. **Genome research**, v. 26, n. 10, p. 1323-1332, 2016.

KADRI, N. K., ZHANG, J., OGET-EBRAD, C., WANG, Y., COULDREY, C., SPELMAN, R., & DRUET, T. High male specific contribution of the X-chromosome to individual global recombination rate in dairy cattle. **BMC genomics**, v. 23, n. 1, p. 114, 2022.

KAUPPI, L., JEFFREYS, A. J., & KEENEY, S. Where the crossovers are: recombination distributions in mammals. **Nature Reviews Genetics**, v. 5, n. 6, p. 413-424, 2004.

KONG, A., BARNARD, J., GUDBJARTSSON, D. F., THORLEIFSSON, G., JONSDOTTIR, G., SIGURDARDOTTIR, S., & STEFANSSON, K. Recombination rate and reproductive success in humans. **Nature genetics**, v. 36, n. 11, p. 1203-1206, 2004.

LOZADA-SOTO, E. A., MALTECCA, C., LU, D., MILLER, S., COLE, J. B., & TIEZZI, F. Trends in genetic diversity and the effect of inbreeding in American Angus cattle under genomic selection. **Genetics Selection Evolution**, v. 53, n. 1, p. 50, 2021.

LU, S., ZONG, C., FAN, W., YANG, M., LI, J., CHAPMAN, A. R., BAI, F. Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. **Science**, v. 338, n. 6114, p. 1627-1630, 2012.

MA L., O'CONNELL J. R., VANRADEN P. M., SHEN B., PADHI A., SUN C. & DA Y. Cattle sex-specific recombination and genetic control from a large pedigree analysis. **PLoS Genetics**, 11:e1005387, 2015.

MALTECCA, C., TIEZZI, F., COLE, J. B., & BAES, C. Symposium review: Exploiting homozygosity in the era of genomics—Selection, inbreeding, and mating programs. **Journal of dairy science**, v. 103, n. 6, p. 5302-5313, 2020.

MEUWISSEN, T. H. E., HAYES, B. J., & GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, n. 4, p. 1819-1829, 2001.

MOURESAN, E. F., GONZÁLEZ-RODRÍGUEZ, A., CAÑAS-ÁLVAREZ, J. J., MUNILLA LEGUIZAMON, S. E. B. A. S. T. I. A. N., ALTARRIBA, J., DIAZ, C. & VARONA, L. Mapping recombination rate on the autosomal chromosomes based on the persistency of linkage disequilibrium phase among autochthonous beef cattle populations in Spain. **Frontiers in Genetics**, v. 10, p. 1170, 2019.

MÜLLER, D., SCHOPP, P., & MELCHINGER, A. E. Selection on expected maximum haploid breeding values can increase genetic gain in recurrent

genomic selection. **G3: Genes, Genomes, Genetics**, v. 8, n. 4, p. 1173-1181, 2018.

NACHMAN, M. W. Variation in recombination rate across the genome: evidence and implications. **Current opinion in genetics & development**, v. 12, n. 6, p. 657-663, 2002.

NIEHOFF, T. A., TEN NAPEL, J., BIJMA, P., POOK, T., WIENTJES, Y. C., HEGEDŰS, B., & CALUS, M. P. Improving selection decisions with mating information by accounting for Mendelian sampling variances looking two generations ahead. **Genetics Selection Evolution**, v. 56, n. 1, p. 41, 2024a.

NIEHOFF, T. A., TEN NAPEL, J., & CALUS, M. P. Prediction of additive genetic variances of descendants for complex families based on Mendelian sampling variances. **G3: Genes, Genomes, Genetics**, p. jkae205, 2024b.

OLLIVIERZ, L. Genetic differences in recombination frequency in the pig (Sus scrofa). **Genome**, v. 38, n. 5, p. 1048-1051, 1995.

PAIGEN, K. & PETKOV, P. Mammalian recombination hot spots: properties, control and evolution. **Nature Reviews Genetics**, v. 11, n. 3, p. 221-233, 2010.
PAIGEN, K. & PETKOV, P. PRDM9 and its role in genetic recombination. **Trends in Genetics**, v. 34, n. 4, p. 291-300, 2018.

PAN, J., SASAKI, M., KNIEWEL, R., MURAKAMI, H., BLITZBLAU, H. G., TISCHFIELD, S. E. & HOCHWAGEN, A. A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. **Cell**, v. 144, n. 5, p. 719-731, 2011.

PEÑALBA, J. V., & WOLF, J. B. From molecules to populations: appreciating and estimating recombination rate variation. **Nature Reviews Genetics**, v. 21, n. 8, p. 476-492, 2020.

PETERSON, A. L., & PAYSEUR, B. A. Sex-specific variation in the genome-wide recombination rate. **Genetics**, v. 217, n. 1, p. iyaa019, 2021.

QANBARI, SABER; WITTENBURG, DÖRTE. Male recombination map of the autosomal genome in German Holstein. **Genetics Selection Evolution**, v. 52, n. 1, p. 73, 2020.

SANDOR, C., LI, W., COPPIETERS, W., DRUET, T., CHARLIER, C., & GEORGES, M. Genetic variants in REC8, RNF212, and PRDM9 influence male recombination in cattle. **PLoS Genet**, v. 8, n. 7, p. e1002854, 2012.

SANTOS, D. J. A., COLE, J. B., LAWLOR JR, T. J., VANRADEN, P. M., TONHATI, H., & MA, L. Variance of gametic diversity and its application in selection programs. **Journal of dairy science**, v. 102, n. 6, p. 5279-5294, 2019.

SANTOS, D. J. A., COLE, J. B., LIU, G. E., VANRADEN, P. M., & MA, L. Gamevar. f90: a software package for calculating individual gametic diversity. **BMC bioinformatics**, v. 21, n. 1, p. 1-5, 2020.

SARDELL, J. M., & KIRKPATRICK, M. Sex differences in the recombination landscape. **The American Naturalist**, v. 195, n. 2, p. 361-379, 2020.

SCHAEFFER, L. R. Strategy for applying genome-wide selection in dairy cattle. **Journal of animal Breeding and genetics**, v. 123, n. 4, p. 218-223, 2006.

SEGELKE, D., REINHARDT, F., LIU, Z., & THALLER, G. Prediction of expected genetic variation within groups of offspring for innovative mating schemes. **Genetics Selection Evolution**, v. 46, n. 1, p. 42, 2014.

SEROUSSI, E., SHIRAK, A., GERSHONI, M., EZRA, E., DE ABREU SANTOS, D. J., MA, L., & LIU, G. E. Bos taurus–indicus hybridization correlates with intralocus sexual-conflict effects of PRDM9 on male and female fertility in Holstein cattle. **BMC genetics**, v. 20, p. 1-11, 2019.

SHEN B., JIANG J., SEROUSSI E., LIU G. E. & MA L. Characterization of recombination features and the genetic basis in multiple cattle breeds. **BMC Genomics**, 19:304, 2018.

TABET, Joe Menwer. Selecting for Mendelian Sampling term to increase variability in a breeding program. 2022.

TAO, S., WU, J., YAO, D., CHEN, Y., YANG, W., & TONG, C. Identification of recombination events in outbred species with next-generation sequencing data. **BMC genomics**, v. 19, n. 1, p. 1-11, 2018.

USDA. Livestock and poultry: world markets and trade. United States Department of Agriculture and Foreign Agricultural Service 31 (2023).

VANRADEN, P. M. Symposium review: How to implement genomic selection. *Journal of Dairy Science*, 2020, 103.6: 5291-5301.

WANG, J., FAN, H. C., BEHR, B., & QUAKE, S. R. Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. **Cell**, v. 150, n. 2, p. 402-412, 2012.

WENG, Z. Q., SAATCHI, M., SCHNABEL, R. D., TAYLOR, J. F., & GARRICK, D. J. Recombination locations and rates in beef cattle assessed from parent-offspring pairs. **Genetics Selection Evolution**, v. 46, n. 1, p. 1-14, 2014.

WENG, Z., WOLC, A., SU, H., FERNANDO, R. L., DEKKERS, J. C., ARANGO, J., & GARRICK, D. J. Identification of recombination hotspots and quantitative trait loci for recombination rate in layer chickens. **Journal of animal science and biotechnology**, v. 10, n. 1, p. 1-13, 2019.

ZHANG, J., KADRI, N. K., MULLAART, E., SPELMAN, R., FRITZ, S., BOICHARD, D., & DRUET, T. Genetic architecture of individual variation in recombination rate on the X chromosome in cattle. **Heredity**, v. 125, n. 5, p. 304-316, 2020.

ZHOU, Y., SHEN, B., JIANG, J., PADHI, A., PARK, K. E., OSWALT, A., ... & MA, L. Construction of PRDM9 allele-specific recombination maps in cattle using large-scale pedigree analysis and genome-wide single sperm genomics. **DNA Research**, v. 25, n. 2, p. 183-194, 2018.

1.3 OBJECTIVES

1.3.1 General objective

The general aim of this thesis is to study meiotic recombination patterns in Nellore cattle and the genetic influence on these events, assembling the first recombination map for genomic data of this breed, and apply this information to estimate the variance of gametic diversity and related parameters for economically important traits. Finally, with the results generated through the previous steps, the last part of the objective is to compare the response to the selection of economically important traits considering the variance of gametic diversity combined with breeding value information and traditional selection by breeding values only.

1.3.2 Specific objectives

- To estimate recombination rates between SNP markers in Nellore
- To identify hotspots regions on Nellore genome
- To estimate heritability and repeatability for recombination rate and their hotspots
- To prospect genomic regions associated with recombination rate and hotspots
- To assembly a valid and useful recombination map for gametic variance estimation
- To estimate SNP marker effects for all traits
- To estimate the variance of gametic diversity for all genotyped animals
- To estimate the coefficient of relative variation
- To apply a selection index for a Nellore cattle population, considering the variance of gametic diversity combined with breeding values information and to verify the impact on the response to selection of including parameters related to the variance of gametic diversity as selection criterion.

CHAPTER 2 - Characterization of recombination events and genome-wide association in Nellore males

ABSTRACT - Understanding genetic diversity in the genome is important for characterizing recombination patterns. Recombination is a fundamental biological process in diploid organisms where parental genetic material is exchanged during meiosis. The study aimed to calculate recombination rates, identify hotspot regions, construct a recombination map for Nellore males, estimate genetic parameters, perform a genomic-wide association study (GWAS) for recombination events, and functionally analyze candidate genes near significant SNPs. Pedigree information comprised 141,153 animals and the conflict.f90 software was used to correct Mendelian errors and impute missing SNP using parental genotypes. We utilized 62,022 animals with imputed genomic information for a high-density panel (409,029 SNPs) based on the ARS-UCD1.2 Bos taurus genome assembly. Recombination rates were assessed using an indirect method based on pedigree. First, the progeny-sire pairs were identified on the pedigree file, and then the genotyped pairs were phased to infer recombination events for paternal meiosis. Hotspot regions were defined as SNP intervals with recombination rate > 2.5 standard deviations above the mean. The genetic parameters for the number of observed crossovers and hotspots per sire were estimated using two models: 1) averaging the events and 2) considering the information as repeated measurements, given that the sires had multiple offspring. Subsequently, the GWAS analysis was performed using both models to identify the genomic regions that significantly affected the recombination rate. The major regions were used for gene annotation, and their functional classification was evaluated. We extracted 407,828 crossover events and 12,357 paternal meiosis, covering 33 Morgans for Nellore males. A total of 2,039 hotspot regions were found, with chromosome 1 presenting the highest number of such regions. The heritability estimates for recombination rate were 0.24 ± 0.06 and 0.17 ± 0.09 for models 1 and 2, respectively, and the repeatability was 0.50 ± 0.03. For the number of hotspots, the heritability was 0.30 ± 0.09 and 0.25 ± 0.06 for models 1 and 2, respectively, and the repeatability was 0.37 ± 0.04. In GWAS analysis, we found significant SNPs for both models (1 - 29 SNPs; 2 - 192 SNPs) and shared genes that may impact recombination directly participating in the homologous recombination process or indirectly influencing the cell cycle, genomic stability, and DNA repair processes. GWAS results for hotspot regions identified the *PRDM9* gene on chromosome 1 in both methods, which is directly related to meiotic recombination. These results provide new insights into the patterns of recombination events that influence the Nellore genome, providing a vital resource for targeted breeding strategies. By developing the first recombination map for Nellore, our research provides a foundation for creating more precise genetic improvement strategies. These strategies aim to enhance productivity, sustainability, and the genetic resilience of Nellore cattle populations.

Keywords: crossover, GWAS, heritability, hotspots, pedigree, phasing, repeatability

2.1 INTRODUCTION

Maintaining genetic diversity is a crucial factor in ensuring the enduring viability of breeding initiatives applied to livestock populations. To comprehend genetic diversity across the Holstein cattle genome in the US, Ma et al. (2015) emphasized the significance of characterizing recombination rate patterns. In eukaryotic sexual species, meiotic recombination patterns are a fundamental biological phenomenon, facilitating the exchange of genetic material between parental entities during the formation of eggs or sperm (BARTON & CHARLESWORTH, 1998; COOP & MYERS, 2007; PAIGEN & PETKOV, 2010). This biological mechanism generates novel combinations of genetic variants, resulting in offspring exhibiting traits distinct from those of their progenitors, thereby contributing to the enhancement of overall genetic diversity. (KAUPPI et al., 2004; KONG et al., 2004; SHEN et al., 2018; JOHNSTON, 2024).

Recombination rates vary widely between and within chromosomes, species, sex and possibly age, that is, this variation is under both genetic and environmental controls (OLLIVIERZ, 1995; BROMAN et al., 1998; SARDELL & KIRKPATRICK, 2020; PEÑALBA et al., 2020; FURMAN et al., 2020; PETERSON & PAYSEUR, 2021). In addition, some studies have suggested that maternal age also had effect on crossover frequency (SHERMAN et al., 1994; WANG et al., 2021; SHEN et al., 2021). Research conducted across various mammalian species, such as mice, dogs, humans, swine, and Atlantic salmon by Dietrich et al. (1996), Neff et al. (1999), Tortereau et al. (2012), Kong et al. (2010), Sardell & Kirkpatrick (2020), Weng et al. (2019), Brekke et al. (2022) and Brekke et al. (2023a), demonstrated higher recombination rates in females compared to males. Ma et al. (2015) evaluated a large sample of Holstein cattle from the USDA national database and, for the first time, reported that bulls exhibit higher levels of recombination than cows. Later studies in cattle showed the same trend, with males showing higher overall recombination rates than females, a pattern especially pronounced in the sub-telomeric regions of autosomes (SHEN et al., 2018; ZHOU et al., 2018; QANBARI & WITTENBURG, 2020; BREKKE et al., 2023b)

In most species, recombination events are not placed randomly throughout the chromosomal DNA but tend to occur in concentrated areas with elevated recombination activity, called hotspot regions (KAUPPI et al., 2004; SHEN et al., 2018; BERGERO et al., 2021). Usually, hotspots are small regions in which recombination rate is higher than in surrounding regions (PAIGEN & PETKOV, 2010; WENG et al., 2014). Considering that recombination is susceptible to genetic and environmental influences, it becomes feasible to manipulate it through environmental modifications or via genetic editing contingent upon the existence of advantageous alleles (BISHOP & EENNNAAM, 2020). Incorporating such interventions into breeding programs has the potential to enhance selection response, thereby increasing genetic diversity in each successive generation (BATTAGIN et al., 2016; BISHOP & EENNNAAM, 2020). An approach that investigated and quantified the potential impact of altering recombination hotspots in genetic improvement programs for livestock was conducted by Gonen et al. (2017). The study's findings revealed that manipulating the positions of these recombination hotspots enabled an increase in the overall selection response and a reduction in genetic variance loss. In addition, higher recombination rates increase the likelihood of beneficial alleles to establish and hinder the establishment of deleterious alleles within a lineage (HARTFIELD & OTTO, 2011).

In cattle, as in other mammals, the *PRDM9* gene plays a central role in determining the location of recombination hotspots. *PRDM9* encodes a protein with a zinc finger domain that recognizes specific DNA sequences, directing where recombination events will occur (PAIGEN & PETKOV, 2018). Several studies in cattle have identified the importance of *PRDM9* in regulating hotspots. For example, Sandor et al. (2012) found that different alleles of *PRDM9* were associated with variation in hotspot locations in cattle. Subsequent studies, such as those by Ahlawat et al. (2017), Zhou et al. (2018), and Seroussi et al. (2019), confirmed the function of *PRDM9* in hotspot diversification in cattle, showing that variations in this gene directly influence recombination patterns within populations. Furthermore, Ma et al. (2015), through a genome-wide association study (GWAS), identified *PRDM9* as one of the key genes responsible for hotspot regulation in cattle, directly linking this gene to the occurrence of recombination hotspots. The variation in the number of zinc finger repeats in *PRDM9* has been

associated with differences in hotspot distribution, especially in sub-telomeric regions, which has implications for genetic diversity in cattle.

Recombination events can be detected using direct and indirect methodologies. The direct methods measure the recombination frequency by examining proteins related to some specific stages of meiosis, such as Single-Sperm Genome-Wide Genotyping (LU et al., 2012; WANG et al., 2012; WANG et al., 2021) or Molecular Assay (PAN et al., 2011). Indirect methods build recombination maps based on the identification of local patterns in Linkage Disequilibrium (MOURESAN et al., 2019) or, as in this study, using phasing genomic information and Pedigree data (MA et al., 2015; SHEN et al., 2018). Indirect methods allow for the estimation of recombination rates at the population level, making them more efficient for estimating gametic variance compared to direct methods. Zhou et al. (2018) validated the reliability of pedigree-based approaches by comparing recombination patterns between sperm and live-born offspring, finding similar patterns across the genome.

Despite the importance of understanding the genetic architecture, inheritance, and recombination patterns of a species or breed for genetic improvement, no specific previous studies have been conducted on meiotic recombination in Nellore cattle or other Zebu breeds. The objectives of this study were to estimate the recombination rate, identify hotspot regions, and construct a recombination map for Nellore males using an indirect method. Additionally, two different models were employed to estimate genetic parameters and perform a genome-wide association study (GWAS) for recombination events. Functional analyses were conducted to identify candidate genes located near significant SNPs identified in both models.

2.2 MATERIAL AND METHODS

The Nellore data used in this study belonged to three commercial breeding programs (DeltaGen, Paint and Cia de Melhoramento), which are part of Alliance Nellore database (www.gensys.com.br). The animals, born between 1984 and 2019, were from 276 commercial herds widely distributed in the Midwest, Southeast, and Northeast of Brazil with high connectedness by the common sires intensively used through artificial insemination (AI), with more than 50% of the

calves born from AI. The animal procedures realized in this research agreed with Animal Care of the São Paulo State University (UNESP), School of Agricultural and Veterinary Science Ethical Committee (protocol number 18.340/16).

2.2.1 Pedigree and genotypic data information

Pedigree information comprised 141,153 animals, with an average inbreeding of 0.18%. A total of 12,984 animals have an inbreeding coefficient of up to 6.25%, while 7 animals exceed an inbreeding coefficient of 25%. Some animals have an unknown parentage, with 41.4% of sires resulting from multiple-sire matings and 23.7% of dams coming from in part from the base population. However, only known pedigree of genotyped sire-progeny pairs was specifically used to estimate the recombination rates. Genotype inconsistencies between parents and progeny were adjusted using conflict.f90 software, which corrects for Mendelian errors and fills missing SNP using parental genotypes (VANRADEN et al., 2015).

A total of 62,022 Nellore animals, 24,042 females and 37,980 males, were genotyped with different Bead chip assay densities aligned to the ARS-UCD1.2 *Bos taurus* genome assembly (Table 1). SNP markers found in the autosomal regions with a GeneCall score exceeding 0.80 were included in the analyses to ensure genotyping quality. The genotyped animals with the different density panels were imputed to the HD panel using FImpute v3 (SARGOLZAEI et al., 2014) considering a reference population of 5,483 genotyped animals, a previous imputation study observed an accuracy higher than 0.97 (CARVALHEIRO et al., 2014).

The qcf90 software (MASUDA et al., 2019) was used for the quality control (QC) of SNP markers and samples. The QC filters in imputed genotypes removed autosomal markers with a minor allele frequency (MAF) lower than 0.05, monomorphic SNPs, and a call rate of markers and samples lower than 0.98. A threshold of 0.16 was applied to exclude SNP markers that did not meet Hardy-Weinberg equilibrium, determined by the maximum deviation of heterozygote frequency from the expected values. After QC filtering, 62,022 genotype animals and 409,029 SNPs remained in the genomic data set for further analyses.

Table 1 – Number of animals and SNP markers before quality control of Nellore animals genotyped in the study

Number of males	Number of females	SNP number	Chip	Manufacturer
721	307	17,794	Zoetis ZL4 (Kalamazoo, MI, USA)	Zoetis
145	-	19,720	GeneSeek® Genomic Profiler 20K – Indicine (Lincoln, NE, USA)	Neogen
49	-	26,151	GGP Bovine LDv3 (Lincoln, NE, USA)	Neogen
6,617	4,210	27,553	Zchip (Araçatuba, SP, BR)	Deoxi
11,703	2,878	29,842	Zoetis Custom SNP ZL5 (Kalamazoo, MI, USA)	Zoetis
1	54	30,108	GGP Bovine LDv3 (Lincoln, NE, USA)	Neogen
10,526	9,334	35,339	GGP Indicus 35K (Lincoln, NE, USA)	Neogen
870	429	54,609	Illumina BovineSNP50 (San Diego, CA, USA)	Illumina
1,569	2,198	54,791	GGP Indicus 50K (Lincoln, NE, USA)	Neogen
3,556	1,372	74,153	GeneSeek® Genomic Profiler 75K – Indicine (Lincoln, NE, USA)	Neogen
2,223	3,260	777,962	Illumina BovineHD BeadChip (San Diego, CA, USA)	Illumina

2.2.2 Analysis and identification of recombination events

The detection of recombination events was carried out using a combination of phasing genotyping and pedigree methods (MA et al., 2015; SHEN et al., 2018). Progeny-sire pairs were extracted from the pedigree of Nellore cattle, both genotyped. A sire may appear in one or multiple pairs because sires show different numbers of progenies. The genotype of each animal was phased based on the sire genotype, and crossover locations were identified by comparing one haplotype of the offspring to the two haplotypes of the corresponding parent (sire). The whole process of finding crossovers includes two steps: 1) phasing and 2) finding crossovers.

Only sires were utilized for the phasing process, which focused on heterozygous SNPs, also known as informative SNPs. For a heterozygote SNP (Aa) in the target individual, the phasing aimed to identify the parental origins of alleles "A" and "a", based on sire's genotype. For example, if the sire has "AA", then allele "A" was inherited from the sire, and allele "a" comes from the dam. On the other hand, if the sire has the genotype "aa", the parental origin can also be determined. After completing the phasing step, crossovers are identified by comparing the phased haplotypes to the sire's haplotypes. The number of meiosis corresponds to the number of progeny-sire pairs, while the total number of crossovers is determined between two informative SNPs (MA et al., 2015).

Recombination maps were created by estimating the recombination rates between consecutive SNPs. Each crossover was equally divided into all SNP intervals between the starting and ending informative SNPs (MA et al., 2015). We analyzed the whole dataset to calculate the number of crossovers between two SNPs. For the QC filter, we followed the specification by Shen et al. (2018), excluding sire-progeny pairs that had more than 60 crossover events across the genome, and all SNP crossovers adjacent to each other were excluded; only SNPs more than ten positions apart were kept for subsequent analyses.

Hotspot regions were defined as SNP intervals with a recombination rate higher than 2.5 standard deviations above the mean, in agreement with previous recombination studies in cattle (SANDOR et al., 2012; MA et al., 215; SHEN et al., 2018). The term “hotspot region” was used instead of “hotspot”, because we used a marker panel of only a few genomic variants in genomic regions rather than sequencing them, and so our SNP intervals were much larger than typical human or mouse recombination hotspots, i.e. a few kb or less (SHEN et al., 2018). Hotspot usage was calculated as the proportion of crossover events that occurred in the hotspot regions per total of meiosis (SHEN et al., 2018).

2.2.3 Genetic parameters

The number of crossover events per meiosis for the sires (second generation) was estimated from each three-generation family and used as phenotypic information. Genetic variance components and heritability estimates

for recombination rates and the number of hotspots were obtained using a single-trait animal model through Bayesian inference. Two models were run: M_1) which considered the number of crossovers observed and the number of hotspots per sire, adjusted for the number of progenies, and M_2) which treated the information as repeated measurements:

$$M_1: \boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{Z_1 a} + \boldsymbol{e}$$

or

$$M_2: \boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{Z_1 a} + \boldsymbol{Z_2 pe} + \boldsymbol{e}$$

where $\boldsymbol{y}$ is the vector information, $\boldsymbol{\mu}$ is the vector of the overall mean, $\boldsymbol{a}$ is the vector of additive genetic effect, $\boldsymbol{pe}$ is the vector of permanent environmental effect. The $\boldsymbol{Z_1}$ and $\boldsymbol{Z_2}$ are the incidence matrix relating $\boldsymbol{y}$ to additive genetic effect ($\boldsymbol{a}$) and permanent environmental effects ($\boldsymbol{pe}$), respectively. The random effects were assumed to be normally distributed as $\boldsymbol{a} \sim N(0, \boldsymbol{G}\sigma_a^2)$, $\boldsymbol{pe} \sim N(0, \boldsymbol{I}\sigma_{pe}^2)$ and $\boldsymbol{e} \sim N(0, \boldsymbol{I}\sigma_e^2)$, where σ_a^2, σ_{pe}^2 and σ_e^2 are variances for additive, permanent environmental and residual effects, respectively, $\boldsymbol{I}$ is an identity matrix and G is the genomic relationship matrix built according to Van Raden (2008) as $\mathrm{G} = \frac{\mathbf{MM'}}{\sum_{i=1}^{w} 2p_i(1-p_i)}$, where $\mathbf{M}$ is the genotype matrix coded 0, 1, and 2 for aa, Aa and AA, and adjusted for allele frequency expressed as $2p_i$, and p_i is the frequency of the second allele calculated from genotypes.

The model was implemented using gibbsf90 software from the blup90 suite (MISZTAL et al., 2018). The Gibbs sampler consisted of a chain of 500,000 cycles, with a burn-in period of the first 50,000 iterations. Samples were stored every five cycles, resulting in the estimation of posterior means for genetic parameters from 90,000 samples. Convergence was evaluated through visual inspection of the trace plot using the BOA package in R (SMITH, 2007) and showed convergence (p-value > 0.25) according to the Geweke test (GEWEKW, 1992).

2.2.4 Genome-wide association analyses

The GWAS for recombination events and the number of hotspots were performed using GCTA v1.93.2 (YANG et al., 2011) with a single-marker linear-mixed-model-based approach. The GWAS analysis was performed considering two models (M1 and M2), and each SNP marker was tested for its association with recombination rate and a number of hotspots; for repeated information, a variable residual variance was considered considering lower residual variance for sires with more measurements. The model used for estimating the SNP marker association can be represented as:

$$M_1: \boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{Wb} + \boldsymbol{Z_1 a} + \boldsymbol{e}$$

or

$$M_2: \boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{Wb} + \boldsymbol{Z_1 a} + \boldsymbol{Z_2 pe} + \boldsymbol{e}$$

where ***W*** is the fixed additive effect of the SNP marker tested, ***b*** is the fixed effect of the SNP marker. Statistical tests for marker effect were performed, standardizing its effect as: $t_k = \frac{\hat{b}_k}{se(\hat{b}_k)}$, where t_k is the t-value for the SNP marker effect; $\hat{b}_k$ is the SNP effects on the phenotypic information and $se(\hat{b}_k)$ is the standard error related to the SNP effect ($\hat{b}_k$). The p-values for the SNP effects were computed as $p - value = 2(1 - \Phi(|t_k|))$, where $\Phi(|t_k|)$ is the cumulative function of the normal distribution for the t-value.

The percentage of genetic variance explained by SNP markers ($\sigma^2_{\hat{u}t}$) was calculated as follows: $\sigma^2_{\hat{u}t} = \frac{Var(\boldsymbol{Z_j}\hat{u}_j)}{\sigma^2_a}\ x\ 100\ \%$, where σ^2_a is the additive genetic variance, $\boldsymbol{Z_j}$ is the vector of the *jth* SNP marker and $\hat{u}_j$ represents the SNP effect of the *jth* SNP marker.

2.2.5 Functional and enrichment analyses

The SNP markers deemed significant were used to search the genes harboring within 100 kb down and upstream of each significant SNP marker using the BioMart tool, from the ENSEMBL software (https://www.ensembl.org/info/docs/tools/vep/index.html), considering the ARS-UCD1.3 *Bos taurus* genome assembly (ROSEN et al., 2020).

The Database for Annotation, Visualization and Integrated Discovery (DAVID; (DAVID Bioinformatics Resources 6.8, http://david.abcc.ncifcrf.gov/) (HUANG et al, 2009) was queried to identify GO classes enriched (biological processes - bp, molecular function - mf, cellular component - cc) and pathways (Kegg), considering the *Bos taurus* as background, for the gene list from GWAS.

2.3 RESULTS AND DISCUSSION

2.3.1 Analysis and identification of recombination events

The methodology employed to detect recombination events in this study focuses on tracking the transmission of alleles from one generation to the next. A total of 29,630 progeny-sire pairs were initially genotyped, leading to the identification of over 3.6 million crossover events. After applying stringent quality control measures, 407,828 crossover events and 12,357 paternal meiosis were retained for analysis. The average number of crossovers per meiosis in Nellore males was calculated to span 33 Morgans, which is notably higher than those reported in other studies involving dairy cattle. For instance, Ma et al. (2015) and Shen et al. (2018) found recombination rates ranging from 22.2 to 25.5 Morgans in Holstein, Jersey, Brown Swiss, and Ayrshire breeds. Similarly, studies by Kadri et al. (2016) and Sandor et al. (2012) reported lower crossover rates in Holstein and Fleckvieh breeds, respectively, reinforcing the observation that recombination rates in Nellore are comparatively higher.

The observed differences in recombination rates between Nellore cattle and dairy breeds can be attributed to a combination of biological, environmental, and methodological factors. First, the structure of the data and the genotyping panels employed play a critical role in the accuracy and resolution of crossover detection, as variations in marker density and coverage can lead to disparities in the estimation of recombination events (PORTO-NETO et al., 2014). Furthermore, the tropical environment where Nellore cattle are predominantly raised introduces unique environmental pressures, such as heat stress and periodic nutritional challenges, which are known to influence genetic and reproductive processes (DEL CORVO et al., 2021). These factors may affect

meiotic mechanisms, potentially altering recombination rates compared to cattle raised in temperate climates.

Additionally, the genomic architecture of Nellore cattle reflects its adaptation to tropical environments, shaped by selective pressures for traits such as heat tolerance, resistance to parasites, and efficient use of low-quality forage. These selective forces have likely contributed to distinct recombination patterns compared to taurine breeds, which have evolved under different environmental and management conditions (FERNANDES et al., 2020; FERNANDES et al., 2022). This evolutionary divergence highlights the impact of both natural and artificial selection in shaping genetic processes, providing a plausible explanation for the unique recombination behaviors observed in Nellore cattle.

Although this study did not allow for a comparison of male and female recombination maps, Ma et al. (2015) reported for the first time in cattle that the male recombination map was 10% longer than the female map, which contrasts with observations in many mammalian species (DIETRICH et al. 1996; NEFF et al., 1999; KONG et al. (2010); TORTEREAU et al., 2012; SARDELL & KIRKPATRICK (2020); WENG et al. (2019); BREKKE et al. (2022) and BREKKE et al. 2023a), that females have more recombination than males. Some theories on the evolution of recombination rates showed that the intense artificial selection directed toward males could be a contributing factor, especially if the selection has a direct or indirect positive effect on the recombination (BLIRT & BELL, 1987; BUTLIN, 2005).

The relative physical position for each SNP interval on a chromosome was calculated by standardizing the original physical position by the chromosome length: a value of zero corresponds to the beginning of a chromosome, and a value of one corresponds to the end. The smooth spline model was fitted across all 29 autosomes (Figure 1; Figure S1). All the cattle autosomes are acrocentric, with centromeres at the beginning and telomeres at the end of chromosomes (BAND et al., 2000). Males had a considerably higher recombination rate near the end of chromosomes (telomeres in cattle) and a lower recombination rate at the beginning (centromeres in cattle). The lower recombination rate in regions close to the centromere is clearly expected since, in these regions, it is well known that the crossing-over event is very low or even that there is no crossing-over in the centromere (SHI et al., 2010; SANTOS et al., 2020).

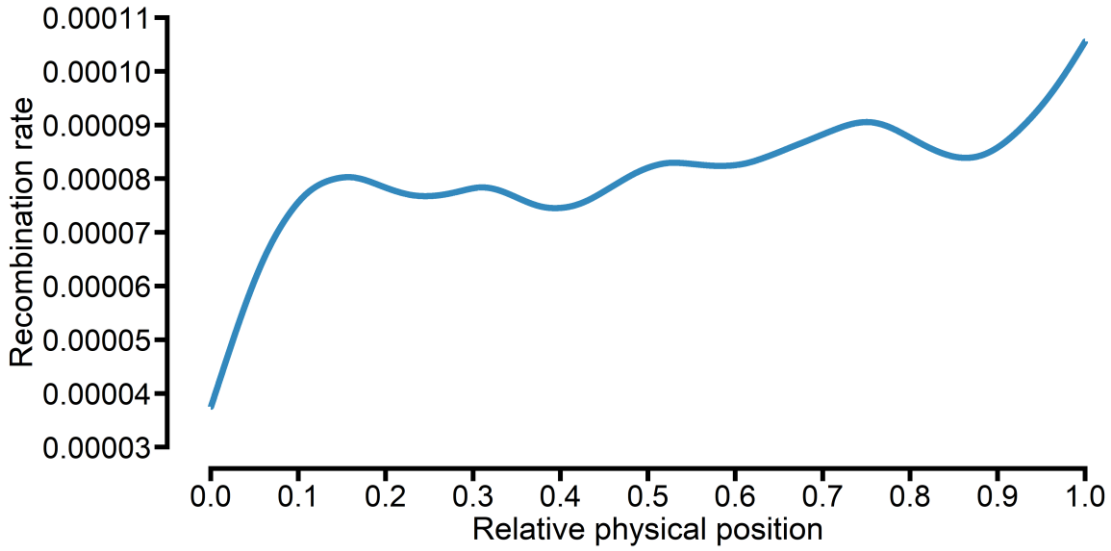


Figure 1 - Distribution of male recombination rates along 29 chromosomes in Nellore cattle

The higher recombination rate found for males at the end of the chromosome is consistent with the observations in previous studies in mice (LYNN et al., 2005), humans (KONG et al., 2010), and cattle (MA et al., 2015; SHEN et al., 2018). The recombination rates are higher in the sub-telomeric regions of chromosomes, which structural and functional factors could explain. The ends of chromosomes have less dense chromatin, making them more accessible to the recombination machinery, thus facilitating crossing-over events (BARTON & CHARLESWORTH, 1998). In contrast, the centromeric regions are more compact and contain a higher density of heterochromatin, which reduces the likelihood of recombination in these areas (PAIGEN & PETKOV, 2010). The concentration of recombination hotspots in the sub-telomeric regions is another key factor contributing to the higher recombination rates observed in these areas (Shi et al., 2010). Studies indicate that hotspots occur more frequently at the ends of chromosomes, while the centromeric regions show lower recombinational activity (PAIGEN & PETKOV, 2018).

Moreover, the *PRDM9* gene, which is known for directing the positioning of recombination events, tends to guide them toward sub-telomeric regions. *PRDM9* activity in these hotspots is associated with the formation of double-strand DNA breaks, which initiates the recombination process (AHLAWAT et al.,

2017). This tendency is observed across many species, including cattle, and has been linked to the preservation of essential genes located in the more central regions of chromosomes, where recombination is less frequent (BARTON & CHARLESWORTH, 1998). From an evolutionary perspective, this organization helps generate genetic variability in regions less prone to deleterious mutations, thereby preserving overall genomic integrity (DIETRICH et al., 1996).

Hotspots were defined as SNP intervals where recombination rates exceed 2.5 and are greater than 30 standard deviations above the average (Figure 2). This definition led to the identification of 2,038 hotspot regions for males (Table S1). While the highest recombination rates were identified on chromosomes 2, 3, 14, and 18 (Figure 3). Notably, chromosomes 1, 2, 3, and 7 had a greater number of hotspot regions, which may be attributed to their longer lengths. Previous studies in cattle have highlighted the variability of recombination rates across different chromosomes. For instance, research by Sandor et al. (2012) and Ma et al. (2015) identified that recombination rates are higher in the telomeric regions of chromosomes, while centromeric regions exhibit lower rates. These studies further showed that certain chromosomes, such as 6, 8, and 21, had a higher density of recombination hotspots in different cattle populations. Kadri et al. (2016), working with Holstein cattle, observed that chromosomes 1, 3, and 5 had elevated recombination rates near the ends of the chromosomes. Our findings align with previous results, show that chromosomes 1, 2, 3, and 7 exhibit a greater number of hotspots, possibly because their extended length allows for increased recombination events. The pattern of recombination hotspot distribution in our study aligns with previous reports, supporting the idea that recombination often tends to concentrate in specific regions, mainly on telomeric regions, across multiple cattle breeds.

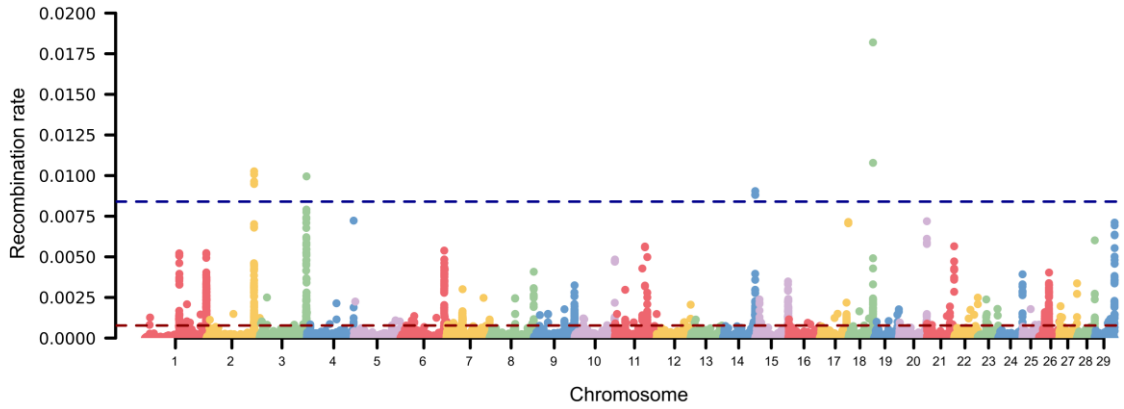


Figure 2- Manhattan plot of recombination rate hotspots that are 2.5 (red dashed line) and 30 (blue dashed line) standard deviations above the mean

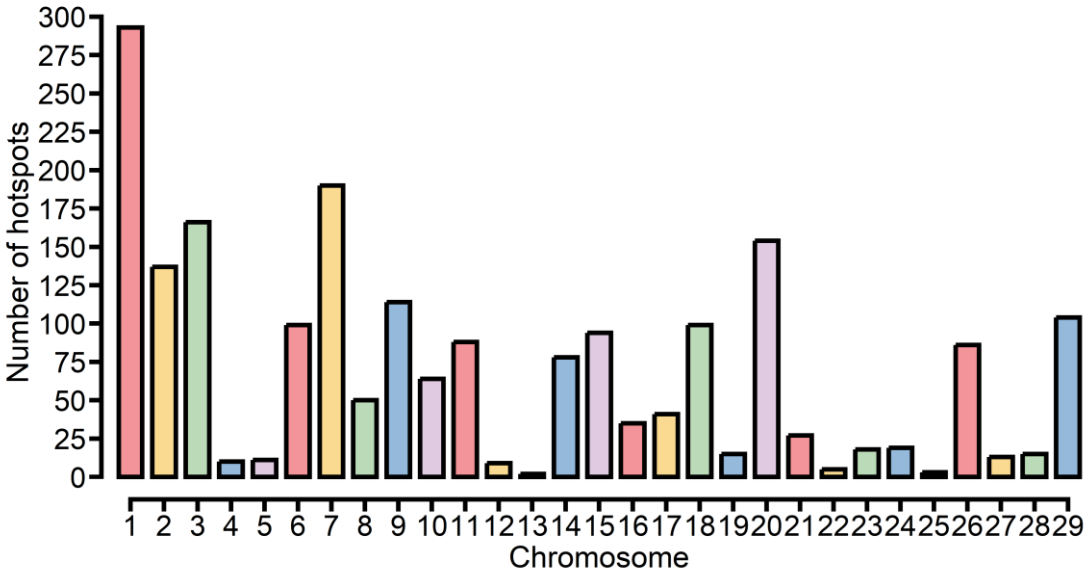


Figure 3 - Distribution of the number of hotspots region with recombination rate > 2.5 standard deviations above the mean per chromosome

After all procedures and analysis described above, the first recombination map for Nellore cattle using Illumina BovineHD BeadChip is available on the supplement of this thesis (Table S2). This map is useful for analysis, such as the estimation of the variance of gametic diversity, and for calculating the genetic distance between genomic regions by an approximation based on physical distance (bp).

2.3.2 Genetic parameters

The heritability estimates for the recombination rate suggest a moderate genetic influence, with values of 0.24 ± 0.06 and 0.17 ± 0.09 for models M1 and M2, respectively, and for the number of hotspots, heritability values of 0.30 ± 0.09 and 0.25 ± 0.06 were observed for models M1 and M2, respectively (Table 2). These values indicate that while genetic factors contribute significantly to the trait, a considerable portion of the variation is also due to non-genetic factors, particularly in model 2, where the permanent environment effect was considered. In line with these findings, previous genomic studies in domestic mammals have consistently reported heritability estimates for recombination rates. Weng et al. (2014) reported heritability estimates of 0.26 ± 0.030 for Angus sires and 0.23 ± 0.042 for Limousin sires, which slightly differ from the results obtained in model M1. Similarly, Sandor et al. (2012) and Johnston et al. (2016), respectively, reported heritability estimates of 0.15 in sheep and 0.22 in cattle for genome-wide recombination rates. However, no heritability estimates for the number of hotspots have been reported in the literature.

The moderate heritability estimates observed suggest that genetic factors significantly influence recombination rates, which can respond to selection, making them effective targets for breeding programs. However, it is important to note that variability in recombination rates is affected by factors such as chromosome length and individual genetic architecture. Additionally, the identification of crossover events might be influenced by parental heterozygosity, which can influence the accuracy of recombination rate estimation (BATTAGIN et al., 2016). Given this complexity, the results from model 2 are likely to provide more robust estimates for the trait, as they better reflect the influence of both genetic and non-genetic factors.

The repeatability estimate obtained from model 2 was 0.50 ± 0.03 for the recombination rate and 0.37 ± 0.04 for the number of hotspots (Table 2), indicating that 50% of the variation in recombination rate and 37% of the variation in the number of hotspots across repeated measurements can be attributed to individual permanent effects, either genetic or environmental. The high repeatability suggests that the recombination rate is a relatively stable trait,

providing confidence in the accuracy of early-life measurements for predicting future performance. When compared to the findings of Weng et al. (2014), who reported repeatability of 0.33 ± 0.027 in Angus sires and 0.30 ± 0.038 in Limousin sires, the higher repeatability observed in our study may be explained by the inclusion of additional covariates in our model, such as bull age, which was found to have a significant effect on recombination rate. Moreover, the adjustment of the phenotype based on the number of informative markers used in genotyping also likely contributed to the higher repeatability observed. However, no repeatability estimate for the number of hotspots has been reported in the literature.

The results highlight the importance of considering both genetic and environmental factors when estimating recombination rates. While genetic factors such as parental heterozygosity and individual chromosome length play a role, the inclusion of covariates such as age and marker information can significantly improve the precision of these estimates. Furthermore, the relatively moderate heritability and high repeatability values suggest that recombination rates have the potential to be included as part of genomic selection strategies aimed at improving genetic diversity and overall breeding efficiency.

Table 2 - Estimates and standard deviation (SD) of direct additive genetic, permanent environmental and residual effect variances, heritability and repeatability for recombination rate obtained in Nellore sires

	Recombination rate		Number of hotspots	
Genetic parameter	*Model 1	*Model 2	*Model 1	*Model 2
Genetic variance	25.29 ± 5.72	10.82 ± 6.08	2.09 ± 0.95	1.99 ± 0.61
Residual variance	82.53 ± 12.19	32.38 ± 6.10	4.68 ± 1.04	5.01 ± 1.15
Permanent environment variance	-	21.97 ± 0.42	-	1.01 ± 0.58
Heritability	0.24 ± 0.06	0.17 ± 0.09	0.30 ± 0.09	0.25 ± 0.06
Repeatability	-	0.50 ± 0.03	-	0.37 ± 0.04

*Model 1) considered the number of crossovers observed and the number of hotspots per sire, adjusted for the number of progenies;
2) treated the information as repeated measurements since the sires had several offspring.

2.3.3 Genome-wide association studies and functional analyses on recombination rate

The GWAS results for recombination rate considering model 1, identified 29 significant SNP markers located on chromosomes 1, 2, 5, 6, 7, 12, 14, 15, 16, and 27 (Figure 4a), which accounted for 3.55% of the additive genetic variance (Table S3). In model 2, a total of 192 SNPs on chromosomes 1, 2, 5, 6, 7, 10, 12, 14, 15, 16, 27, and 29 (Figure 4b) explained 10.97% of the additive genetic variance for recombination rate (Table S3). Despite the differing number of significant SNPs identified using the two models, the genes surrounding these SNPs were shared among the models (Table 3).

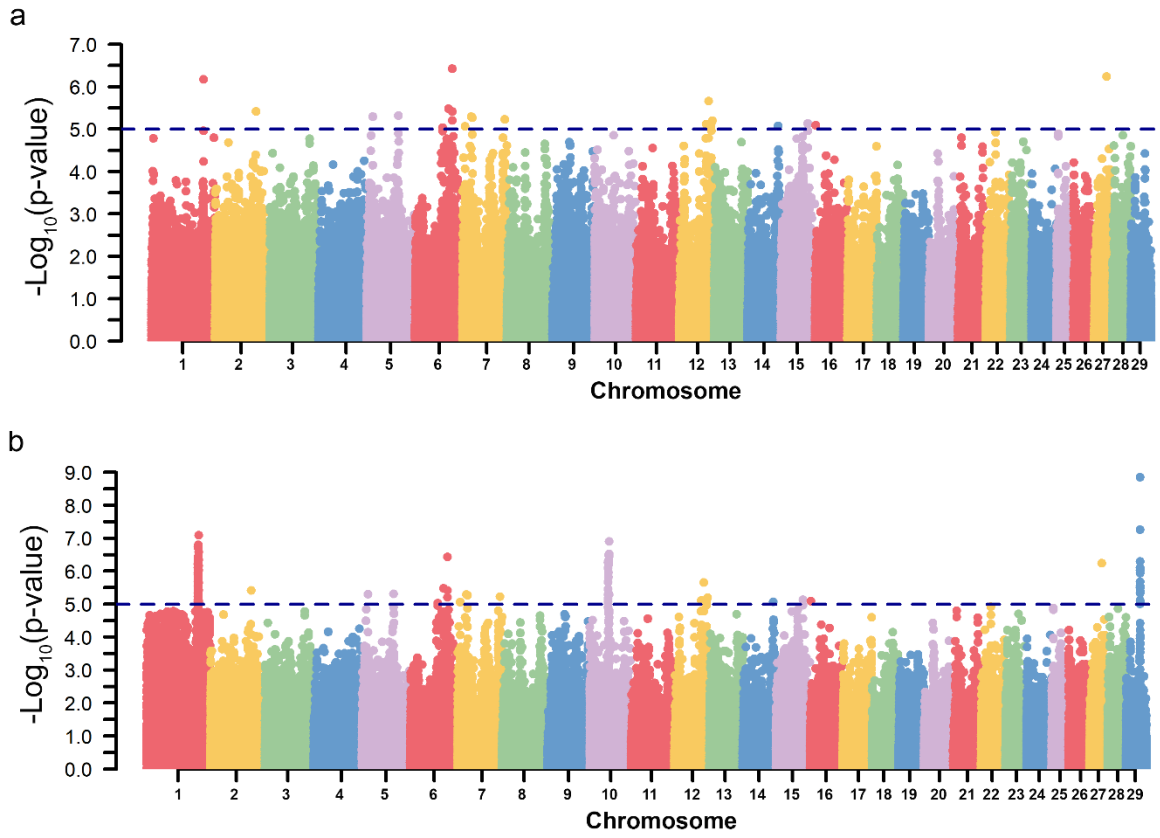


Figure 4 - Manhattan plots of the GWAS of genome-wide recombination rates for Nellore males a) Using average number of crossovers per sire b) Using repeated measures, since the sires had several offspring.

Table 3 – Chromosome (CHR), start and end position in bp and gene symbol and gene name found in the windows close to the significant SNPs

CHR	GENE START	GENE END	GENE SYMBOL	GENE NAME	GENE STABLE ID
1	129434207	129548904	*NMNAT3*	nicotinamide nucleotide adenylyl transferase 3	ENSBTAG00000051936
1	129626685	129651105	*RBP2*	retinol binding protein 2	ENSBTAG00000015056
1	129556625	129583702	*RBP1*	retinol binding protein 1	ENSBTAG00000020028
5	14673266	14877358	*LRRIQ1*	leucine rich repeats and IQ motif containing 1	ENSBTAG00000027900
5	14916523	14939197	*ALX1*	ALX homeobox 1	ENSBTAG00000014977
5	78797983	78841092	*CAPRIN2*	caprin family member 2	ENSBTAG00000019373
5	78855639	78920921	*IPO8*	importin 8	ENSBTAG00000014106
6	92761853	93292876	*FRAS1*	Fraser extracellular matrix complex subunit 1	ENSBTAG00000010716
6	92907305	92907362	*bta-mir-2285bf-1*	bta-mir-2285bf-1	ENSBTAG00000054044
6	83291686	83321882	*STAP1*	signal transducing adaptor family member 1	ENSBTAG00000000962
7	6384274	6395767	*SMIM7*	small integral membrane protein 7	ENSBTAG00000013613
7	6397521	6453129	*MED26*	mediator complex subunit 26	ENSBTAG00000005312
7	6455315	6472363	*SLC35E1*	solute carrier family 35 member E1	ENSBTAG00000005295
7	6479012	6498287	*CHERP*	calcium homeostasis endoplasmic reticulum protein	ENSBTAG00000005289
10	45158053	45165480	*PTGDR*	prostaglandin D2 receptor	ENSBTAG00000006703
10	45268611	45277443	*PIF1*	PIF1 5'-to-3' DNA helicase	ENSBTAG00000007638
10	45311967	45337970	*RBPMS2*	RNA binding protein, mRNA processing factor 2	ENSBTAG00000011037
10	45645738	45653956	*PCLAF*	PCNA clamp associated factor	ENSBTAG00000039462
10	45661323	45816217	*CSNK1G1*	casein kinase 1 gamma 1	ENSBTAG00000016823
10	45818677	45824617	*PPIB*	peptidylprolyl isomerase B	ENSBTAG00000016822
10	45880414	45895810	*CIAO2A*	cytosolic iron-sulfur assembly component 2A	ENSBTAG00000002012
10	45905873	46089594	*DAPK2*	death associated protein kinase 2	ENSBTAG00000011820
10	46185094	46398046	*HERC1*	HECT and RLD domain containing E3 ubiquitin protein ligase family member 1	ENSBTAG00000019474
10	47971129	47971233	*U6*	U6 spliceosomal RNA	ENSBTAG00000042731
12	82802493	82802553	*bta-mir-2301*	bta-mir-2301	ENSBTAG00000053710

12	65755150	67339324	*GPC5*	glypican 5	ENSBTAG00000008708
12	73363796	74075781	*HS6ST3*	heparan sulfate 6-O-sulfotransferase 3	ENSBTAG00000039065
14	75914395	76103864	*CNGB3*	cyclic nucleotide gated channel subunit beta 3	ENSBTAG00000036304
16	2255105	2292885	*MDM4*	MDM4 regulator of p53	ENSBTAG00000006255
27	30732560	31371432	*UNC5D*	unc-5 netrin receptor D	ENSBTAG00000010241

The *HERC1* gene encodes an E3 ubiquitin ligase that is crucial for various cellular processes by regulating the ubiquitination and subsequent degradation of target proteins (PEDRAZZA et al., 2023). This gene plays a pivotal role in the ubiquitin-mediated proteolysis pathway (bta04120), essential for maintaining cellular homeostasis and regulating protein turnover. The ubiquitin-proteasome system (UPS) has been implicated in multiple stages of oocyte meiotic maturation in mice and pigs, including spindle assembly, chromosome alignment, and polar body extrusion, highlighting its critical role in reproductive physiology (HUO, FAN, LIANG, et al., 2004; HUO, FAN, ZHONG, et al., 2004).

HERC1 has also been enriched in biological processes linked to autophagy (GO:0006914, GO:0010507), a conserved intracellular degradation and recycling system that plays an essential role in maintaining cellular homeostasis, regulating cell fate, and supporting reproductive development (GAO et al., 2019). Although autophagy is primarily a cytoplasmic process, deficiencies in this system can lead to genomic instability, increased DNA damage, and impaired cellular repair mechanisms (KARANTZA-WADSWORTH et al., 2007). Notably, inhibition of autophagy has been associated with elevated proteasomal activity, which enhances the degradation of checkpoint kinase 1 (*CHK1*), a crucial protein for homologous recombination during DNA repair (GILLESPIE & RYAN, 2016; LIU et al., 2015).

In bovines, while studies directly investigating *HERC1* are lacking, the UPS has been shown to play a role in fertility by aiding the recycling and remodeling of gamete structures during fertilization. This system is active in both oocytes and spermatozoa, underlining its importance in bovine reproduction (SUTOVSKY, 2018). Furthermore, the UPS and autophagy serve as the primary intracellular quality control mechanisms in eukaryotes, using ubiquitination as a degradation signal. The UPS targets short-lived and misfolded soluble proteins, while autophagy handles long-lived proteins, insoluble aggregates, and entire organelles, ensuring cellular homeostasis (KOCATURK & GOZUACIK, 2018).

Autophagy is also essential for reproductive processes. In males, it plays a significant role in spermatogenesis, supporting events such as testosterone production, the assembly of ectoplasmic specialization, spermatid differentiation, and the maintenance of spermatogonia stem cells (SHANG et al., 2016; WANG et al., 2021; YIN et al., 2017). In females, autophagy is critical during oocyte

maturation; its inhibition significantly impairs developmental potential by disrupting mitochondrial function, triggering DNA damage responses, and inducing early apoptosis, as observed in porcine oocytes matured in vitro (SHEN et al., 2018). These findings underscore the essential role of autophagy in proper oocyte development and subsequent embryonic progression.

Functional analysis has revealed that *MDM4* is a key regulator of the p53 signaling pathway (bta04115). *MDM4* acts as a negative regulator of p53 by binding to its transactivation domain, thus modulating its activity in response to cellular stress (TOLEDO & WAHL, 2007). The p53 protein itself is a central player in numerous cellular processes, including cell growth, differentiation, senescence, and DNA repair, making it a pivotal guardian of genomic stability (VOUSDEN & LU, 2002). In vertebrates, p53 also plays a significant role in maintaining germline genomic integrity and regulating reproductive functions (HU, 2009).

Evidence from murine models underscores the importance of p53 in reproductive biology. Mice with null or reduced levels of p53 exhibit germ cell degeneration during the meiotic prophase within the testicular seminiferous tubules, indicating its crucial role in meiosis (ROTTER et al., 1993). Furthermore, the expression of p53 in spermatocytes suggests its direct involvement in meiosis, particularly in recombination and DNA repair processes (NAKAMUTA & KOBAYASHI, 2004). Recent findings further highlight the essential role of p53 in meiotic recombination, demonstrating that it is required for proper crossover formation and localization during meiosis (MARCET-ORTEGA et al., 2022).

In the context of meiotic recombination, p53 appears to coordinate DNA repair mechanisms to maintain genomic integrity, a function that is particularly critical in germ cells due to their role in transmitting genetic information to the next generation. Through its interaction with *MDM4*, p53 activity is tightly regulated to prevent excessive or insufficient DNA damage responses, ensuring a balance that supports both cellular homeostasis and reproductive efficiency. Although studies directly linking *MDM4* and p53 to bovine reproduction are scarce, the evolutionary conservation of the p53 signaling pathway across vertebrates suggests that similar mechanisms may exist in cattle.

PIF1 is a highly conserved DNA helicase with multiple functions in mitochondria and cell nucleus (SABOURI et al., 2012). It plays a role in DNA

synthesis in difficult-to-replicate regions and in lagging-strand synthesis in conjunction with polymerase (BUDD et al., 2006; PAESCHKE et al., 2011; SABOURI et al., 2012). Furthermore, this helicase is involved in maintaining the mitochondrial genome, maturing Okazaki fragments, resolving G-quadruplex structures, and participating in DNA synthesis during breakage-induced replication and cross-recombination (BUDD et al., 2006; CHUNG, 2014; PAESCHKE et al., 2011; PIKE et al., 2009; ROSSI et al., 2008; WILSON et al., 2013). Vernekar et al. (2021) reported that PIF1 is actively involved in meiotic recombination and DNA double-strand break (DSB) repair, suggesting a contribution to the normal homologous recombination machinery. Additionally, *PIF1* is also involved in modulating telomerase activity and DNA synthesis during repair processes. Nickens et al. (2018) highlighted the synergistic effects of *PIF1* with other helicases in preparing DNA repair intermediates for homologous recombination and protecting DSBs from telomere additions, as well in the maintenance of telomere length, thus suggesting its participation in ensuring genome integrity.

IPO8, a member of the nuclear transporter family, is essential for the movement of proteins from the cytoplasm to the nucleus by recognizing nuclear targeting signals (WAGSTAFF & JANS, 2009). This gene plays a crucial role in meiosis by facilitating the transport of several proteins involved in key cell cycle processes, including the regulation of chromosome attachment to the nuclear envelope, chromosome movement, meiotic recombination, chromosome segregation, and the timely incorporation of structural proteins necessary for meiotic progression (BOWMAN et al., 2019; PALACIOS et al., 2021). A notable function of *IPO8* is its role in transporting microRNA-containing nuclear complexes (GOURISHETTI et al., 2023), as well as mediating the nuclear import of *SMAD4*, a key component of the TGF-β signaling pathway, which is critical for cell cycle arrest (LARIVIÈRE et al., 2014; YAO et al., 2008). The interplay between *IPO8* and *SMAD4* highlights the importance of nuclear transport in regulating cellular responses to developmental signals.

SMAD4, as a mediator of the TGF-β signaling pathway, plays a vital role in germ cell differentiation and meiotic progression. Studies have shown that *SMAD4* deletion leads to defects in meiosis in female germ cells, underscoring its importance in reproductive biology (WU et al., 2013, 2015, 2016).

Furthermore, *SMAD4* is widely expressed in the testis, particularly in gonocytes, Leydig cells, and Sertoli cells, indicating its involvement in testicular development and spermatogenesis (ZHANG et al., 2011). These cell types are central to the regulation of spermatogenesis, with Leydig cells producing testosterone, Sertoli cells providing structural and nutritional support to germ cells, and gonocytes serving as precursors to mature sperm.

Together, *IPO8* and *SMAD4* form part of an intricate network of molecular interactions essential for meiosis and reproductive development. Their contributions to processes such as meiotic recombination, chromosome segregation, and germ cell differentiation highlight their fundamental roles in maintaining genomic integrity and fertility across sexes. The nuclear transport function of *IPO8* ensures the precise localization of *SMAD4* and other critical factors, facilitating the orchestration of these complex biological pathways.

A cluster of genes (*CAPRIN2*, *GPC5*, and *CSNK1G1*) was enriched for the biological process of positive regulation of the canonical Wnt signaling pathway (GO:0090263). The Wnt signaling pathway plays a crucial role in diverse cellular processes such as proliferation, differentiation, migration, and polarity in various organisms (LV et al., 2023). Evidence has shown that this pathway is active in gonads and regulates spermatogenesis and oogenesis through β-catenin signaling, a key component of the canonical Wnt pathway (LI et al., 2014). Activation of the Wnt pathway has also been demonstrated to stimulate the resumption of meiosis in human primordial germ cells, underscoring its critical role in germline development and gametogenesis (YANG et al., 2022). In rainbow trout, Wnt pathway-related genes were found to be expressed during gonadal differentiation and gametogenesis, emphasizing the evolutionary importance of this pathway in reproductive processes (NICOL & GUIGUEN, 2012).

In bovines, the Wnt signaling pathway plays an essential role in follicular development and the selection of the dominant follicle. Gupta et al. (2014) demonstrated that genes involved in the Wnt pathway, such as *CTNNB1* (β-catenin), *FZD6*, and *AXIN2*, are dynamically regulated during follicular waves, a characteristic process in monotocous species like cattle. The canonical Wnt pathway enhances follicular responsiveness to FSH by upregulating key genes, including *CYP19A1* (aromatase), which is critical for estradiol production, and *CCND2*, which promotes granulosa cell proliferation. This suggests a functional

link between the Wnt pathway and follicular growth, with *CAPRIN2*, *GPC5*, and *CSNK1G1* potentially contributing to this regulation.

Specifically, *CAPRIN2* may interact with β-catenin to influence transcriptional activity, while *GPC5*, as a glypican family member, modulates Wnt ligand-receptor interactions, facilitating pathway activation. *CSNK1G1*, a casein kinase, participates in the phosphorylation of β-catenin and other pathway components, ensuring precise regulation of Wnt signaling during follicular development. Gupta et al. (2014) also noted that β-catenin expression is higher in dominant follicles compared to subordinate ones, emphasizing the importance of Wnt signaling in follicular selection and dominance.

This pathway's ability to coordinate cell proliferation, differentiation, and hormonal responsiveness highlights its central role in the regulation of reproductive efficiency in bovines. The interaction between Wnt signaling and FSH-driven follicular development provides a framework for understanding how genes like *CAPRIN2*, *GPC5*, and *CSNK1G1* contribute to reproductive success. These findings underscore the significance of *CAPRIN2*, *GPC5*, and *CSNK1G1* in the regulation of follicular development, linking their roles to key processes such as follicular dynamics, meiotic progression, and gametogenesis in cattle. Their involvement in the Wnt signaling pathway highlights the intricate molecular mechanisms underlying reproductive efficiency in bovines.

Functional enrichment analysis identified *MED26*, *ALX1*, and *PPIB* as genes involved in the regulation of transcription by RNA polymerase II (GO:0045944, GO:0006357) and RNA polymerase binding (GO:0070063). These processes are intricately connected to a phenomenon known as transcription-associated recombination (TAR). TAR is a well-documented process across various organisms, from prokaryotes to higher eukaryotes, and occurs when DNA replication and transcription elongation proceed simultaneously (GOTTIPATI et al., 2008; AYGÜN & SVEJSTRUP, 2010).

TAR arises when collisions between transcription and replication machinery obstruct the replication fork, resulting in increased recombination frequencies as a mechanism to resolve these conflicts and maintain genomic stability (PRADO & AGUILERA, 2005). During transcription, newly synthesized RNA molecules can hybridize with the DNA template to form R-loops—structures that interfere with replication fork progression. These stalled forks are often

resolved through recombination pathways, underscoring the functional relationship between transcription and recombination (AGUILERA, 2002). Evidence further suggests that transcription increases a target region's accessibility to recombination machinery and DNA damage responses, facilitating the recruitment of recombination proteins and promoting recombination occurrence (ALT et al., 1986; GARCÍA-RUBIO et al., 2003). In mammals, TAR is notably involved in developmentally regulated processes such as immunoglobulin gene class switching, a mechanism critical for antibody diversification, which is both replication- and recombination-dependent (CHAUDHURI & ALT, 2004; LI et al., 2004; GOTTIPATI et al., 2008).

The genes *MED26*, *ALX1*, and *PPIB* may play crucial roles in transcription-recombination dynamics. As a component of the Mediator complex, *MED26* bridges interactions between transcriptional activators and RNA polymerase II, facilitating the formation of preinitiation complexes essential for efficient transcription. Its involvement suggests a potential role in coordinating transcriptional regulation with recombination processes. *ALX1* is a transcription factor known for its role in developmental processes and DNA binding. *ALX1* could influence the accessibility of specific genomic regions to recombination proteins by modulating chromatin structure during transcription. *PPIB e*ncoding cyclophilin B, *PPIB* participates in protein folding and may influence transcription indirectly by stabilizing transcriptional machinery or modifying protein-protein interactions critical for RNA polymerase II activity.

In mammals, TAR not only resolves replication-transcription conflicts to maintain genomic stability but also serves as a mechanism for generating genetic diversity, critical for adaptive processes such as immune responses. The association of *MED26*, *ALX1*, and *PPIB* with transcription regulation and RNA polymerase II binding underscores their potential significance in facilitating these processes and highlights their relevance in recombination-related phenomena, particularly in the context of maintaining genomic integrity and promoting genetic diversity.

Among the non-coding RNAs identified in this study, the *microRNAs bta-mir-2285bf-1* and *bta-mir-2301* stand out as potential regulators of critical biological processes in cattle. MicroRNAs (miRNAs) are key post-transcriptional regulators that influence gene expression by binding to target mRNAs,

modulating their stability and translation. Studies have indicated that *bta-mir-2285bf* is involved in the regulation of early embryonic development, showing a significant negative correlation with blastocyst rates, suggesting that this miRNA may directly or indirectly affect the expression of genes essential for early embryonic development (SOUZA, 2019). On the other hand, *bta-mir-2301* has been identified in miRNA profiling analyses of mammary tissues in dairy cows, although its specific functions remain largely unexplored (LI et al., 2018). These miRNAs exemplify the intricate regulatory networks mediated by non-coding RNAs and their importance in shaping genetic and phenotypic diversity in cattle.

2.3.4 Identification of *PRDM9* through GWAS Targeting Meiotic Recombination Hotspots

The GWAS results for recombination hotspots considering model a, identified 44 significant SNP markers located on chromosomes 1, 3, 10, 11, 12, 13, 15, and 26 (Figure 5a), which accounted for 1.29% of the additive genetic variance (Table S4). In model b, a total of 52 SNPs on chromosomes 1, 3, 4, 6, 8, 9, 10, 12, 16, 26, and 28 (Figure 5b) explained 1.32% of the additive genetic variance for recombination hotspots (Table S4). Despite the differing number of significant SNPs identified using the two models, the genes surrounding these SNPs were shared among the models (Table 4).

The GWAS results on recombination hotspots provide further detailed insights into the genetic factors regulating recombination hotspots in cattle, highlighting the identification of the *PRDM9* gene as a locus strongly associated with the presence of recombination hotspots in both models (Figure 5a and 5b). A significant association of *PRDM9* is observed on chromosome 1, confirming its central role in determining the location of hotspots, consistent with previous studies in cattle and other species. Additionally, other relevant genomic regions were identified (Model a: Chromosomes 1, 3, 10, 11, 12, 13, 15, and 26; Model b: Chromosomes 1, 3, 4, 6, 8, 9, 10, 12, 16, 26, and 28), suggesting that multiple loci influence variation in recombination rate. These findings reinforce the complexity of the molecular mechanisms controlling recombination, with *PRDM9* being the main regulator but with the participation of other genes that may

modulate hotspot activity and distribution throughout the genome. Although both models identified the *PRDM9* gene, the results also highlight significant differences between the analyzed models (Figures 5a and 5b), indicating that the genetic contribution to recombination may vary depending on the parameters and methods used, which expands our understanding of the genetic architecture of this complex trait.

The hotspots are essential for shaping the genomic landscape and influencing patterns of inheritance within cattle populations (KAUPPI et al., 2004; PAIGEN & PETKOV, 2010). The *PRDM9* gene is a key regulator of recombination hotspot placement, which plays an essential role in determining the location of these events during meiosis. *PRDM9*, through its zinc finger domain, binds to specific DNA sequences, directing where DNA breaks will occur to facilitate genetic exchange between homologous chromosomes (PAIGEN & PETKOV, 2018). In cattle, variation in *PRDM9* sequences, especially in the number of zinc finger repeats, contributes to differences in hotspot patterns and distribution, which directly impacts genetic diversity (ZHOU et al., 2018). This diversity is fundamental for the evolution of cattle populations and for selective breeding processes, as it enhances genetic variability (AHLAWAT et al., 2017; SEROUSSI et al., 2019).

In addition to its role in hotspot regulation, *PRDM9* also influences fertility and speciation in cattle. Studies have shown that specific *PRDM9* variants are associated with reduced recombination rates in hybrids, particularly in crosses between Bos taurus and Bos indicus, where incompatibilities between *PRDM9* alleles can lead to male infertility (SEROUSSI et al., 2019). This phenomenon mirrors what has been observed in other species, such as mice, where *PRDM9* is the only known gene associated with hybrid sterility (AHLAWAT et al., 2017; PAIGEN & PETKOV, 2018). In cattle, the identification of different *PRDM9* variants has highlighted significant differences in hotspot usage between alleles, which directly affects recombination rates, and the genetic diversity passed on to offspring (SANDOR et al., 2012; ZHOU et al., 2018). Understanding these dynamics is critical for improving breeding strategies and managing genetic diversity in cattle populations.

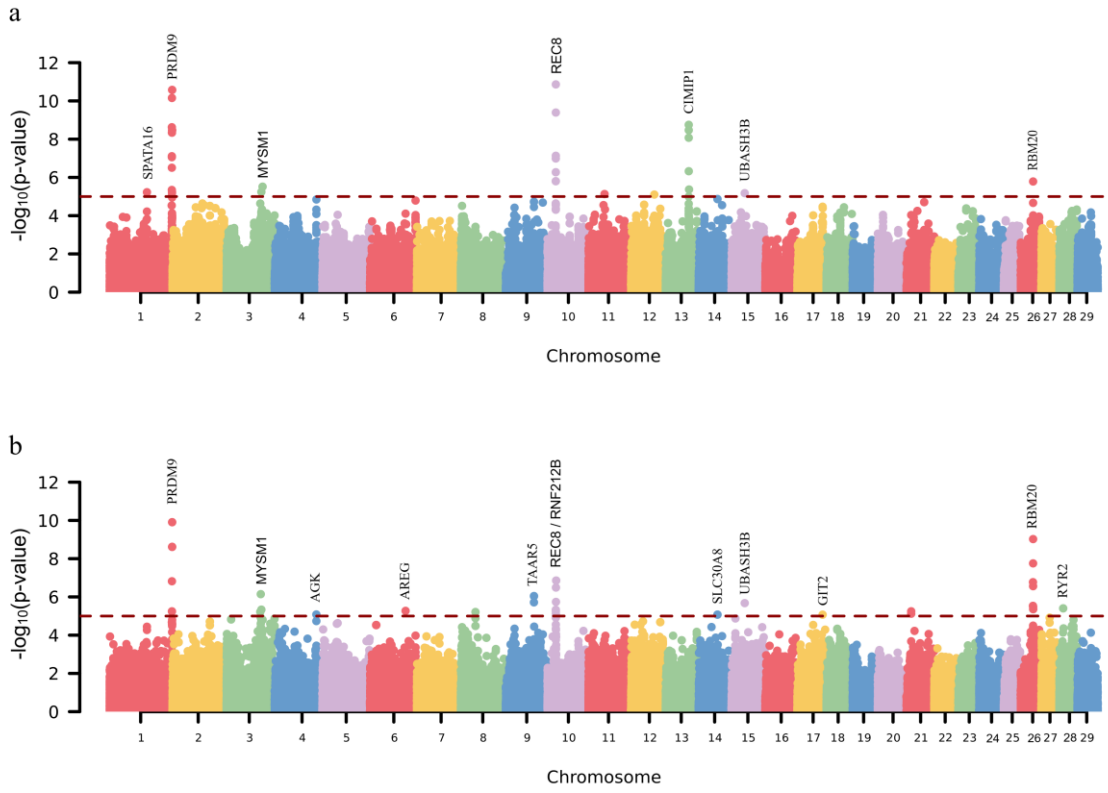


Figure 5 - Manhattan plots of the GWAS of genome-wide hotspot usage for Nellore males a) Using average number of crossovers per sire b) Using repeated measures, since the sires had several offspring

Table 4 – Chromosome (CHR), start and end position in bp and gene symbol and gene name found in the windows close to the significant SNPs

CHR	GENE START	GENE END	GENE SYMBOL	GENE NAME	GENE STABLE ID
1	94320212	94597998	*SPATA16*	spermatosis associated 16	ENSBTAG00000031802
1	156340298	156842780	*KCNH8*	potassium voltage-gated channel subfamily H member 8	ENSBTAG00000012798
1	157125017	157190483	*EFHB*	EF-hand domain family member B	ENSBTAG00000013806
1	157217981	157232069	*RAB5A*	RAB5A, member RAS onco family	ENSBTAG00000046385
1	157232251	157249458	*PP2D1*	protein phosphatase 2C like domain containing 1	ENSBTAG00000004061
1	157277360	157376183	*KAT2B*	lysine acetyltransferase 2B	ENSBTAG00000000746
1	157392691	157411259	*SGO1*	shugoshin 1	ENSBTAG00000000752
1	157489488	157505082	*ZNF596*	zinc finger protein 596	ENSBTAG00000049158
1	157545666	157559397	*PRDM9*	PR domain containing 9	ENSBTAG00000004538
3	87349526	87391325	*MYSM1*	Myb like, SWIRM and MPN domains 1	ENSBTAG00000013961
3	87455279	87459484	*TACSTD2*	tumor associated calcium signal transducer 2	ENSBTAG00000004381
3	87479311	87544530	*OMA1*	OMA1 zinc metallopeptidase	ENSBTAG00000017326
3	84203465	84816448	*NFIA*	nuclear factor I A	ENSBTAG00000000074
4	104316388	104729340	*TMEM178B*	transmembrane protein 178B	ENSBTAG00000051724
4	104801940	104902617	*AGK*	acylglycerol kinase	ENSBTAG00000019265
4	104913725	104960329	*DENND11*	DENN domain containing 11	ENSBTAG00000016879
6	89306919	89325427	*EREG*	epiregulin	ENSBTAG00000010273
6	89379645	89391792	*AREG*	amphiregulin	ENSBTAG00000018134
9	70753522	70754535	*TAAR5*	trace amine associated receptor 5	ENSBTAG00000047985
9	70784406	70785326	*TAAR2*	trace amine associated receptor 2	ENSBTAG00000038414
9	70806940	70807950	*TAAR1*	trace amine associated receptor 1	ENSBTAG00000037718
9	70834660	70857677	*VNN1*	vanin 1	ENSBTAG00000015094
9	70846744	70846841	*U6*	U6 spliceosomal RNA	ENSBTAG00000042998
10	20855686	20860122	*LTB4R2*	leukotriene B4 receptor 2	ENSBTAG00000040056
10	20858358	20865354	*CIDEB*	cell death inducing DFFA like effector b	ENSBTAG00000010236

10	20865333	20871269	*NOP9*	NOP9 nucleolar protein	ENSBTAG00000010235
10	20871434	20880388	*DHRS1*	dehydrogenase/reductase 1	ENSBTAG00000010234
10	20893284	20900213	*RABGGTA*	Rab geranylgeranyltransferase subunit alpha	ENSBTAG00000018796
10	20901907	20916712	*TGM1*	transglutaminase 1	ENSBTAG00000003920
10	20926605	20936522	*TINF2*	TERF1 interacting nuclear factor 2	ENSBTAG00000002718
10	20927368	20936187	*GMPR2*	guanosine monophosphate reductase 2	ENSBTAG00000002715
10	20948208	20957579	*CHMP4A*	charged multivesicular body protein 4A	ENSBTAG00000039415
10	20949085	20955298	*TSSK4*	testis specific serine kinase 4	ENSBTAG00000002710
10	20962178	20967158	*TM9SF1*	transmembrane 9 superfamily member 1	ENSBTAG00000002709
10	20965617	20965669	*bta-mir-11989*	bta-mir-11989	ENSBTAG00000050750
10	20967440	20978204	*IPO4*	importin 4	ENSBTAG00000002707
10	20976598	20982410	*REC8*	REC8 meiotic recombination protein	ENSBTAG00000002705
10	20986949	20991549	*IRF9*	interferon regulatory factor 9	ENSBTAG00000005816
10	20992427	21003698	*RNF31*	ring finger protein 31	ENSBTAG00000005815
10	21004530	21007790	*PSME2*	proteasome activator subunit 2	ENSBTAG00000005814
10	21010402	21014209	*EMC9*	ER membrane protein complex subunit 9	ENSBTAG00000025778
10	21012734	21016071	*PSME1*	proteasome activator subunit 1	ENSBTAG00000021395
10	21017366	21019157	*FITM1*	fat storage inducing transmembrane protein 1	ENSBTAG00000021394
10	21021441	21029337	*DCAF11*	DDB1 and CUL4 associated factor 11	ENSBTAG00000021392
10	21035549	21044740	*PCK2*	phosphoenolpyruvate carboxykinase 2, mitochondrial	ENSBTAG00000011934
10	21055767	21060433	*NRL*	neural retina leucine zipper	ENSBTAG00000001871
10	21062390	21069364	*CPNE6*	copine 6	ENSBTAG00000001870
10	21070624	21087501	*CARMIL3*	capping protein regulator and myosin 1 linker 3	ENSBTAG00000015208
10	21088245	21100627	*DHRS4*	dehydrogenase/reductase 4	ENSBTAG00000017665
10	21088245	21100627	*DHRS4*	dehydrogenase/reductase 4	ENSBTAG00000017665
10	21017366	21019157	*FITM1*	fat storage inducing transmembrane protein 1	ENSBTAG00000021394
10	21371184	21380016	*JPH4*	junctophilin 4	ENSBTAG00000007104
10	21380687	21389478	*AP1G2*	adaptor related protein complex 1 subunit gamma 2	ENSBTAG00000007099
10	21388435	21392402	*THTPA*	thiamine triphosphatase	ENSBTAG00000007097
10	21405142	21424338	*ZFHX2*	zinc finger homeobox 2	ENSBTAG00000007053

10	21461588	21468762	*NGDN*	neuroguidin	ENSBTAG00000009595
10	21490410	21510517	*MYH7*	myosin heavy chain 7	ENSBTAG00000009703
10	21505501	21505608	*MIR208B*	microRNA 208b	ENSBTAG00000037339
10	21515534	21542213	*MYH6*	myosin heavy chain 6	ENSBTAG00000040053
10	21534412	21534494	*MIR208A*	microRNA 208a	ENSBTAG00000029863
10	21544234	21546707	*CMTM5*	CKLF like MARVEL transmembrane domain containing 5	ENSBTAG00000032709
10	21547535	21551904	*IL25*	interleukin 25	ENSBTAG00000009701
10	21559539	21569078	*EFS*	embryonal Fyn-associated substrate	ENSBTAG00000038733
10	21572374	21578551	*SLC22A17*	solute carrier family 22 member 17	ENSBTAG00000016467
10	21590471	21597349	*PABPN1*	poly(A) binding protein nuclear 1	ENSBTAG00000006884
10	21602048	21605997	*BCL2L2*	BCL2 like 2	ENSBTAG00000019692
10	21607231	21612728	*PPP1R3E*	protein phosphatase 1 regulatory subunit 3E	ENSBTAG00000037679
10	21613371	21659773	*RNF212B*	ring finger protein 212B	ENSBTAG00000025755
10	21620210	21630418	*HOMEZ*	homeobox and leucine zipper encoding	ENSBTAG00000032686
10	21842083	21858614	*C10H14orf93*	chromosome 10 C14orf93 homolog	ENSBTAG00000012725
10	21862575	21871855	*AJUBA*	ajuba LIM protein	ENSBTAG00000012724
10	21888397	21896197	*HAUS4*	HAUS augmin like complex subunit 4	ENSBTAG00000012723
10	21915919	21924136	*PRMT5*	protein arginine methyltransferase 5	ENSBTAG00000010890
10	21924916	21956955	*RBM23*	RNA binding motif protein 23	ENSBTAG00000038333
10	21941858	21946478	*REM2*	RRAD and GEM like GTPase 2	ENSBTAG00000016381
10	21949027	21955730	*LRP10*	LDL receptor related protein 10	ENSBTAG00000016378
10	21966320	21978292	*MMP14*	matrix metallopeptidase 14	ENSBTAG00000014824
10	21979455	21990995	*MRPL52*	mitochondrial ribosomal protein L52	ENSBTAG00000014822
10	21990496	22029418	*SLC7A7*	solute carrier family 7 member 7	ENSBTAG00000014821
10	22027169	22067494	*OXA1L*	OXA1L mitochondrial inner membrane protein	ENSBTAG00000014820
13	58171994	58187787	*CIMIP1*	ciliary microtubule inner protein 1	ENSBTAG00000016926
13	58889622	58975046	*BMP7*	bone morphotic protein 7	ENSBTAG00000015362
13	58954605	58954673	*bta-mir-2307*	bta-mir-2307	ENSBTAG00000044456
14	47110107	47149980	*SLC30A8*	solute carrier family 30 member 8	ENSBTAG00000052098
15	33291476	33439897	*UBASH3B*	ubiquitin associated and SH3 domain containing B	ENSBTAG00000000842

17	63256628	63262336	*C12orf76*	chromosome 12 open reading frame 76	ENSBTAG00000032047
17	63266230	63300591	*ANKRD13A*	ankyrin repeat domain 13A	ENSBTAG00000014376
17	63304695	63345212	*GIT2*	GIT ArfGAP 2	ENSBTAG00000006506
17	63350533	63365189	*TCHP*	trichoplein keratin filament binding	ENSBTAG00000006504
17	63377041	63399500	*GLTP*	glycolipid transfer protein	ENSBTAG00000002591
17	63410291	63453042	*TRPV4*	transient receptor potential cation channel subfamily V member 4	ENSBTAG00000000031
26	31202434	31396918	*RBM20*	RNA binding motif protein 20	ENSBTAG00000023891
26	31436457	31463087	*PDCD4*	programmed cell death 4	ENSBTAG00000019434
26	31451966	31452046	*bta-mir-6524*	bta-mir-6524	ENSBTAG00000052505
26	31462433	31462508	*bta-mir-4680*	bta-mir-4680	ENSBTAG00000047843
26	31463888	31480690	*BBIP1*	BBSome interacting protein 1	ENSBTAG00000046362
26	31062315	31062382	*bta-mir-2394*	bta-mir-2394	ENSBTAG00000044608
26	31063540	31076435	*DUSP5*	dual specificity phosphatase 5	ENSBTAG00000020283
26	31124276	31162340	*SMC3*	structural maintenance of chromosomes 3	ENSBTAG00000013905
26	31481196	31568116	*SHOC2*	SHOC2 leucine rich repeat scaffold protein	ENSBTAG00000007709
26	31582908	31591909	*ADRA2A*	adrenoceptor alpha 2A	ENSBTAG00000039292
28	9691435	10518660	*RYR2*	ryanodine receptor 2	ENSBTAG00000022886

2.4 CONCLUSION

We characterized the recombination patterns in Nellore males and provided the first recombination map based on 407,828 crossover events and 12,357 paternal meiosis, covering 33 Morgans for Nellore males. We identified 2,038 hotspot regions, with chromosome 1 having the greatest number of hotspot regions. An SNP panel derived from these results is being made available to the scientific community. The heritability estimate suggests that the meiotic recombination events are inherited in Nellore cattle. The GWAS analyses identified significant SNPs and genes related to recombination, some acting directly in the homologous recombination process and others influencing the cell cycle, genomic stability, or DNA repair. Using recombination hotspot regions, the *PRDM9* gene on chromosome 1 was identified, which is directly linked to meiotic recombination. These results provide new insights into meiotic recombination in Nellore cattle, with a focus on male recombination maps and varying recombination rates across different chromosome regions, particularly in sub-telomeric areas. Overall, our findings highlight the value of the genomic resources developed over years of genetic evaluation in Nellore cattle and offer important opportunities to better understand the genetic mechanisms and evolutionary processes governing them.

2.5 REFERENCES

AGUILERA, Andrés. The connection between transcription and genomic instability. **The EMBO journal**, v. 21, n. 3, p. 195-201, 2002.

AHLAWAT, S., DE, S., SHARMA, P., SHARMA, R., ARORA, R., KATARIA, R. S., & SINGH, R. K. Evolutionary dynamics of meiotic recombination hotspots regulator PRDM9 in bovids. **Molecular genetics and genomics**, v. 292, p. 117-131, 2017.

ALT, Frederick W. et al. Regulation of genome rearrangement events during lymphocyte differentiation. **Immunological reviews**, v. 89, p. 5-30, 1986.

AYGÜN, Ozan; SVEJSTRUP, Jesper Q. RECQL5 helicase: connections to DNA recombination and RNA polymerase II transcription. **DNA repair**, v. 9, n. 3, p. 345-353, 2010.

BAND, M. R., LARSON, J. H., REBEIZ, M., GREEN, C. A., HEYEN, D. W., DONOVAN, J., ... & LEWIN, H. A. An ordered comparative map of the cattle and human genomes. **Genome research**, v. 10, n. 9, p. 1359-1368, 2000.

BARTON, N. H.; CHARLESWORTH, B. Why sex and recombination? **Science**, v. 281, n. 5385, p. 1986-1990, 1998.

BATTAGIN, M., GORJANC, G., FAUX, A. M., JOHNSTON, S. E., & HICKEY, J. M. Effect of manipulating recombination rates on response to selection in livestock breeding programs. **Genetics Selection Evolution**, v. 48, p. 1-12, 2016.

BERGERO, R., ELLIS, P., HAERTY, W., LARCOMBE, L., MACAULAY, I., MEHTA, T., & IMMLER, S. Meiosis and beyond–understanding the mechanistic and evolutionary processes shaping the germline genome. **Biological Reviews**, v. 96, n. 3, p. 822-841, 2021.

BISHOP, T. F., & VAN EENENNAAM, A. L. Genome editing approaches to augment livestock breeding programs. **Journal of Experimental Biology**, v. 223, n. Suppl_1, p. jeb207159, 2020.

BOWMAN, R., BALUKOF, N., FORD, T., & SMOLIKOVE, S. (2019). A Novel Role for α-Importins and Akirin in Establishment of Meiotic Sister Chromatid Cohesion in Caenorhabditis elegans. *Genetics*, *211*(2), 617–635.

BREKKE, C., BERG, P., GJUVSLAND, A. B., & JOHNSTON, S. E. Recombination rates in pigs differ between breeds, sexes and individuals, and are associated with the RNF212, SYCP2, PRDM7, MEI1 and MSH4 loci. **Genetics Selection Evolution**, v. 54, n. 1, p. 33, 2022.

BREKKE, C., JOHNSTON, S. E., KNUTSEN, T. M., & BERG, P. Genetic architecture of individual meiotic crossover rate and distribution in Atlantic Salmon. **Scientific Reports**, v. 13, n. 1, p. 20481, 2023a.

BREKKE, C., JOHNSTON, S. E., GJUVSLAND, A. B., & BERG, P. Variation and genetic control of individual recombination rates in Norwegian Red dairy cattle. **Journal of Dairy Science**, v. 106, n. 2, p. 1130-1141, 2023b.

BROMAN, K. W., MURRAY, J. C., SHEFFIELD, V. C., WHITE, R. L., & WEBER, J. L. Comprehensive human genetic maps: individual and sex-specific variation in recombination. **The American Journal of Human Genetics**, v. 63, n. 3, p. 861-869, 1998.

BUDD, M. E., REIS, C. C., SMITH, S., MYUNG, K., & CAMPBELL, J. L. Evidence suggesting that Pif1 helicase functions in DNA replication with the Dna2 helicase/nuclease and DNA polymerase δ. **Molecular and cellular biology**, v. 26, n. 7, p. 2490-2500, 2006.

BURT, A., & BELL, G. Mammalian chiasma frequencies as a test of two theories of recombination. **Nature**, v. 326, n. 6115, p. 803-805, 1987.

BUTLIN, R. K. Recombination and speciation. **Molecular Ecology**, v. 14, n. 9, p. 2621-2635, 2005.

CHAUDHURI, Jayanta; ALT, Frederick W. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. **Nature Reviews Immunology**, v. 4, n. 7, p. 541-552, 2004.

CHUNG, W. H. To peep into Pif1 helicase: multifaceted all the way from genome stability to repair-associated DNA synthesis. **Journal of Microbiology**, v. 52, p. 89-98, 2014.

COOP, G., & MYERS, S. R. Live hot, die young: transmission distortion in recombination hotspots. **PLoS genetics**, v. 3, n. 3, p. e35, 2007.

DEL CORVO, M., LAZZARI, B., CAPRA, E., ZAVAREZ, L., MILANESI, M., UTSUNOMIYA, Y. T., & AJMONE-MARSAN, P. Methylome patterns of cattle adaptation to heat stress. **Frontiers in genetics**, v. 12, p. 633132, 2021.

DIETRICH W. F., MILLER J., STEEN R., MERCHANT M. A., DAMRON-BOLES D., HUSAIN Z., DREDGE R., DALY M. J., INGALLS K. A. & O'CONNOR T.J. A comprehensive genetic map of the mouse genome. **Nature**, 380:149–52, 1996.

FERNANDES JUNIOR, G. A., DE OLIVEIRA, H. N., CARVALHEIRO, R., CARDOSO, D. F., FONSECA, L. F. S., VENTURA, R. V., & DE ALBUQUERQUE, L. G. Whole-genome sequencing provides new insights into genetic mechanisms of tropical adaptation in Nellore (Bos primigenius indicus). **Scientific reports**, v. 10, n. 1, p. 9412, 2020.

FERNANDES JUNIOR, G. A., PERIPOLLI, E., SCHMIDT, P. I., CAMPOS, G. S., MOTA, L. F. M., MERCADANTE, M. E. Z., & DE ALBUQUERQUE, L. G. Current applications and perspectives of genomic selection in Bos indicus (Nellore) cattle. **Livestock Science**, v. 263, p. 105001, 2022.

FURMAN, B. L., METZGER, D. C., DAROLTI, I., WRIGHT, A. E., SANDKAM, B. A., ALMEIDA, P., & MANK, J. E. Sex chromosome evolution: so many exceptions to the rules. **Genome biology and evolution**, v. 12, n. 6, p. 750-763, 2020.

GAO, H., KHAWAR, M. B., & LI, W. Autophagy in reproduction. **Autophagy: Biology and Diseases: Basic Science**, p. 453-468, 2019.

GARCIA-RUBIO, M. et al. Recombinogenic effects of DNA-damaging agents are synergistically increased by transcription in Saccharomyces cerevisiae: new insights into transcription-associated recombination. **Genetics**, v. 165, n. 2, p. 457-466, 2003.

GEWEKE J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernado JM, Berger JO, Smith AP, Dawid AFM, editors. **Bayesian Statistics. 4th ed.** Oxford: Clarendon Press. p. 169–93; 1992.

GILLESPIE, D. A., & RYAN, K. M. Autophagy is critically required for DNA repair by homologous recombination. **Molecular & Cellular Oncology**, v. 3, n. 1, p. e1030538, 2016.

GONEN, S., BATTAGIN, M., JOHNSTON, S. E., GORJANC, G., & HICKEY, J. M. The potential of shifting recombination hotspots to increase genetic gain in livestock breeding. **Genetics Selection Evolution**, v. 49, p. 1-12, 2017.

GOTTIPATI, Ponnari et al. Transcription-associated recombination is dependent on replication in Mammalian cells. **Molecular and cellular biology**, v. 28, n. 1, p. 154-164, 2008.

GOURISHETTI, K., BALAJI EASWARAN, V., MOSTAKIM, Y., RANGANATH PAI, K. S., & BHERE, D. MicroRNA (miR)-124: A Promising Therapeutic Gateway for Oncology. **Biology**, v. 12, n. 7, p. 922, 2023.

GUPTA, P. P., FOLGER, J. K., RAJPUT, S. K., LV, L., YAO, J., IRELAND, J. J., & SMITH, G. W. Regulation and regulatory role of WNT signaling in potentiating FSH action during bovine dominant follicle selection. **PloS one**, v. 9, n. 6, p. e100201, 2014.

HARTFIELD M. & OTTO S. P. Recombination and hitchhiking of deleterious alleles. **Evolution: International Journal of Organic Evolution**, 65:2421- 2434, 2011.

HU, W. (2009). The Role of p53 Gene Family in Reproduction. ***Cold Spring Harbor Perspectives in Biology***, *1*(6), a001073.

HUO, L. J., FAN, H. Y., LIANG, C. G., YU, L. Z., ZHONG, Z. S., CHEN, D. Y., & SUN, Q. Y. Regulation of ubiquitin-proteasome pathway on pig oocyte meiotic maturation and fertilization. **Biology of reproduction**, v. 71, n. 3, p. 853-862, 2004a.

HUO, L. J., FAN, H. Y., ZHONG, Z. S., CHEN, D. Y., SCHATTEN, H., & SUN, Q. Y. Ubiquitin–proteasome pathway modulates mouse oocyte meiotic maturation and fertilization via regulation of MAPK cascade and cyclin B1 degradation. **Mechanisms of development**, v. 121, n. 10, p. 1275-1287, 2004.

JOHNSTON, S. E., BÉRÉNOS, C., SLATE, J., & PEMBERTON, J. M. Conserved genetic architecture underlying individual recombination rate variation in a wild population of Soay sheep (Ovis aries). **Genetics**, v. 203, n. 1, p. 583-598, 2016.

JOHNSTON, SUSAN E. Understanding the genetic basis of variation in meiotic recombination: past, present, and future. **Molecular Biology and Evolution**, v. 41, n. 7, p. msae112, 2024.

KADRI, N. K., HARLAND, C., FAUX, P., CAMBISANO, N., KARIM, L., COPPIETERS, W., ... & DRUET, T. Coding and noncoding variants in HFM1, MLH3, MSH4, MSH5, RNF212, and RNF212B affect recombination rate in cattle. **Genome research**, v. 26, n. 10, p. 1323-1332, 2016.

KARANTZA-WADSWORTH, V., PATEL, S., KRAVCHUK, O., CHEN, G., MATHEW, R., JIN, S., & WHITE, E. Autophagy mitigates metabolic stress and genome damage in mammary tumorigenesis. **Genes & development**, v. 21, n. 13, p. 1621-1635, 2007.

KAUPPI, L., JEFFREYS, A. J., & KEENEY, S. Where the crossovers are: recombination distributions in mammals. **Nature Reviews Genetics**, v. 5, n. 6, p. 413-424, 2004.

KOCATURK, N. M., & GOZUACIK, D. Crosstalk between mammalian autophagy and the ubiquitin-proteasome system. **Frontiers in cell and developmental biology**, v. 6, p. 128, 2018.

KONG, A., BARNARD, J., GUDBJARTSSON, D. F., THORLEIFSSON, G., JONSDOTTIR, G., SIGURDARDOTTIR, S., & STEFANSSON, K. Recombination rate and reproductive success in humans. **Nature genetics**, v. 36, n. 11, p. 1203-1206, 2004.

KONG, A., THORLEIFSSON, G., GUDBJARTSSON, D. F., MASSON, G., SIGURDSSON, A., JONASDOTTIR, A., & STEFANSSON, K. Fine-scale recombination rate differences between sexes, populations and individuals. **Nature**, v. 467, n. 7319, p. 1099-1103, 2010.

LARIVIÈRE, N., LAW, J., & TRINKLE-MULCAHY, L. Dissection of a novel autocrine signaling pathway via quantitative secretome and interactome mapping. **Journal of Proteome Research**, v. 13, n. 7, p. 3432-3443, 2014.

LI, Z., WOO, C. J., IGLESIAS-USSEL, M. D., RONAI, D., & SCHARFF, M. D. The generation of antibody diversity through somatic hypermutation and class switch recombination. **Genes & development**, v. 18, n. 1, p. 1-11, 2004.

LI, H., ZHANG, Z., BI, Y., YANG, D., ZHANG, L., & LIU, J. Expression characteristics of β-catenin in scallop Chlamys farreri gonads and its role as a

potential upstream gene of Dax1 through canonical Wnt signalling pathway regulating the spermatogenesis. **PLoS One**, v. 9, n. 12, p. e115917, 2014.

LI, Q., YANG, C., DU, J., ZHANG, B., HE, Y., HU, Q., ... & ZHONG, J. Characterization of miRNA profiles in the mammary tissue of dairy cattle in response to heat stress. **BMC genomics**, v. 19, p. 1-11, 2018.

LIU, E. Y., XU, N., O'PREY, J., LAO, L. Y., JOSHI, S., LONG, J. S., O'PREY, M., CROFT, D. R., BEAUMATIN, F., BAUDOT, A. D., MRSCHTIK, M., ROSENFELDT, M., ZHANG, Y., GILLESPIE, D. A., & RYAN, K. M. Loss of autophagy causes a synthetic lethal deficiency in DNA repair. **Proceedings of the National Academy of Sciences**, v. 112, n. 3, p. 773-778, 2015.

LU, S., ZONG, C., FAN, W., YANG, M., LI, J., CHAPMAN, A. R., BAI, F. Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. **Science**, v. 338, n. 6114, p. 1627-1630, 2012.

LV, Z., XU, H., SI, X., XU, S., LI, X., LI, N., ZHOU, Q., CHANG, M., YAO, S., & LI, H. XAV-939 inhibits epithelial-mesenchymal transformation in pulmonary fibrosis induced by crystalline silica via the Wnt signaling pathway. **Environmental Toxicology**, v. 38, n. 2, p. 460-471, 2023.

LYNN, A., SCHRUMP, S., CHERRY, J., HASSOLD, T., & HUNT, P. Sex, not genotype, determines recombination levels in mice. **The American Journal of Human Genetics**, v. 77, n. 4, p. 670-675, 2005.

MA L., O'CONNELL J. R., VANRADEN P. M., SHEN B., PADHI A., SUN C. & DA Y. Cattle sex-specific recombination and genetic control from a large pedigree analysis. **PLoS Genetics**, 11:e1005387, 2015.

MARCET-ORTEGA, M., MALDONADO-LINARES, A., LÓPEZ-PANADÉS, M., & ROIG, I. p53 controls meiotic prophase progression and crossover formation. **International journal of molecular sciences**, v. 23, n. 17, p. 9818, 2022.

MISZTAL I, TSURUTA S, LOURENCO D, AGUILAR I, LEGARRA A, VITEZICA Z. **Manual for BLUPF90 family of programs**. Athens: University of Georgia; 2018.

MOURESAN, E. F., GONZÁLEZ-RODRÍGUEZ, A., CAÑAS-ÁLVAREZ, J. J., MUNILLA LEGUIZAMON, S. E. B. A. S. T. I. A. N., ALTARRIBA, J., DIAZ, C. & VARONA, L. Mapping recombination rate on the autosomal chromosomes based on the persistency of linkage disequilibrium phase among autochthonous beef cattle populations in Spain. **Frontiers in Genetics**, v. 10, p. 1170, 2019.

NAKAMUTA, N., & KOBAYASHI, S. Developmental expression of p63 in the mouse testis. **Journal of veterinary medical science**, v. 66, n. 6, p. 681-687, 2004.

NEFF M. W., BROMAN K. W., MELLERSH C. S., RAY K., ACLAND G. M., AGUIRRE G. D., ZIEGLE J. S., OSTRANDER E. A. & RINE J. A second-generation genetic linkage map of the domestic dog, *Canis familiaris*. **Genetics**, 151:803–20, 1999.

NICKENS, D. G., ROGERS, C. M., & BOCHMAN, M. L. The Saccharomyces cerevisiae Hrq1 and Pif1 DNA helicases synergistically modulate telomerase activity in vitro. **Journal of Biological Chemistry**, v. 293, n. 37, p. 14481-14496, 2018.

NICOL, B., & GUIGUEN, Y. Expression profiling of Wnt signaling genes during gonadal differentiation and gametogenesis in rainbow trout. **Sexual development**, v. 5, n. 6, p. 318-329, 2012.

OLLIVIERZ, L. Genetic differences in recombination frequency in the pig (Sus scrofa). **Genome**, v. 38, n. 5, p. 1048-1051, 1995.

PAESCHKE, K., CAPRA, J. A., & ZAKIAN, V. A. DNA replication through G-quadruplex motifs is promoted by the Saccharomyces cerevisiae Pif1 DNA helicase. **Cell**, v. 145, n. 5, p. 678-691, 2011.

PAIGEN, K. & PETKOV, P. Mammalian recombination hotspots: properties, control and evolution. **Nature Reviews Genetics**, v. 11, n. 3, p. 221-233, 2010.

PAIGEN, K. & PETKOV, P. PRDM9 and its role in genetic recombination. **Trends in Genetics**, v. 34, n. 4, p. 291-300, 2018.

PALACIOS, V., KIMBLE, G. C., TOOTLE, T. L., & BUSZCZAK, M. Importin-9 regulates chromosome segregation and packaging in Drosophila germ cells. **Journal of cell science**, v. 134, n. 7, p. jcs258391, 2021.

PAN, J., SASAKI, M., KNIEWEL, R., MURAKAMI, H., BLITZBLAU, H. G., TISCHFIELD, S. E. & HOCHWAGEN, A. A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. **Cell**, v. 144, n. 5, p. 719-731, 2011.

PEDRAZZA, L., MARTINEZ-MARTINEZ, A., SÁNCHEZ-DE-DIEGO, C., VALER, J. A., PIMENTA-LOPES, C., SALA-GASTON, J., & ROSA, J. L. *HERC1* deficiency causes osteopenia through transcriptional program dysregulation during bone remodeling. **Cell Death & Disease**, v. 14, n. 1, p. 17, 2023.

PEÑALBA, J. V., & WOLF, J. B. From molecules to populations: appreciating and estimating recombination rate variation. **Nature Reviews Genetics**, v. 21, n. 8, p. 476-492, 2020.

PETERSON, A. L., & PAYSEUR, B. A. Sex-specific variation in the genome-wide recombination rate. **Genetics**, v. 217, n. 1, p. iyaa019, 2021.

PIKE, J. E., BURGERS, P. M. J., CAMPBELL, J. L., & BAMBARA, R. A. Pif1 helicase lengthens some Okazaki fragment flaps necessitating Dna2 nuclease/helicase action in the two-nuclease processing pathway. **Journal of Biological Chemistry**, v. 284, n. 37, p. 25170-25180, 2009.

PORTO-NETO, L. R., REVERTER, A., PRAYAGA, K. C., CHAN, E. K., JOHNSTON, D. J., HAWKEN, R. J., & BARENDSE, W. The genetic architecture of climatic adaptation of tropical cattle. **Plos one**, v. 9, n. 11, p. e113284, 2014.

PRADO, Félix; AGUILERA, Andrés. Impairment of replication fork progression mediates RNA polII transcription-associated recombination. **The EMBO journal**, v. 24, n. 6, p. 1267-1276, 2005.

QANBARI, SABER; WITTENBURG, DÖRTE. Male recombination map of the autosomal genome in German Holstein. **Genetics Selection Evolution**, v. 52, n. 1, p. 73, 2020.

ROSSI, M. L., PIKE, J. E., WANG, W., BURGERS, P. M. J., CAMPBELL, J. L., & BAMBARA, R. A. Pif1 helicase directs eukaryotic Okazaki fragments toward the two-nuclease cleavage pathway for primer removal. **Journal of Biological Chemistry**, v. 283, n. 41, p. 27483-27493, 2008.

ROTTER, V., SCHWARTZ, D., ALMON, E., GOLDFINGER, N., KAPON, A., MESHORER, A., DONEHOWER, L. A., & LEVINE, A. J. Mice with reduced levels of p53 protein exhibit the testicular giant-cell degenerative syndrome. **Proceedings of the National Academy of Sciences**, v. 90, n. 19, p. 9075-9079, 1993.

SABOURI, N., MCDONALD, K. R., WEBB, C. J., CRISTEA, I. M., & ZAKIAN, V. A. DNA replication through hard-to-replicate sites, including both highly transcribed RNA Pol II and Pol III genes, requires the S. pombe Pfh1 helicase. **Genes & development**, v. 26, n. 6, p. 581-593, 2012.

SANDOR, C., LI, W., COPPIETERS, W., DRUET, T., CHARLIER, C., & GEORGES, M. Genetic variants in REC8, RNF212, and PRDM9 influence male recombination in cattle. **PLoS Genet**, v. 8, n. 7, p. e1002854, 2012.

SANTOS, D. J. A.; CAMARGO, G. M. F., CARDOSO, D. F.; BUZANSKAS, M. E.; ASPILCUETA-BORQUIS, R. R.; HURTADO-LUGO, N. A.; ARAÚJO NETO, F.

A.; ALBUQUERQUE, L. G.; MA, L.; TONHATI, H. Linkage Disequilibrium-Based Inference of Genome Homology and Chromosomal Rearrangements Between Species. **G3: Gene, Genomics, Genetics**, 10:7, 2020.

SARDELL, J. M., & KIRKPATRICK, M. Sex differences in the recombination landscape. **The American Naturalist**, v. 195, n. 2, p. 361-379, 2020.

SEROUSSI, E., SHIRAK, A., GERSHONI, M., EZRA, E., DE ABREU SANTOS, D. J., MA, L., & LIU, G. E. Bos taurus–indicus hybridization correlates with intralocus sexual-conflict effects of PRDM9 on male and female fertility in Holstein cattle. **BMC genetics**, v. 20, p. 1-11, 2019.

SHANG, Y., WANG, H., JIA, P., ZHAO, H., LIU, C., LIU, W., SONG, Z., XU, Z., YANG, L., WANG, Y., & LI, W. Autophagy regulates spermatid differentiation via degradation of PDLIM1. **Autophagy**, v. 12, n. 9, p. 1575-1592, 2016.

SHEN B., JIANG J., SEROUSSI E., LIU G. E. & MA L. Characterization of recombination features and the genetic basis in multiple cattle breeds. **BMC Genomics**, 19:304, 2018.

SHEN, X. H., JIN, Y. X., LIANG, S., KWON, J. W., ZHU, J. W., LEI, L., & KIM, N. H. Autophagy is required for proper meiosis of porcine oocytes maturing in vitro. **Scientific Reports**, v. 8, n. 1, p. 12581, 2018.

SHEN, B., FREEBERN, E., JIANG, J., MALTECCA, C., COLE, J. B., LIU, G. E., & MA, L. Effect of temperature and maternal age on recombination rate in cattle. **Frontiers in Genetics**, v. 12, p. 682718, 2021.

SHERMAN, S. L., PETERSEN, M. B., FREEMAN, S. B., HERSEY, J., PETTAY, D., TAFT, L., & HASSOLD, T. J. Non-disjunction of chromosome 21 in maternal meiosis I: evidence for a maternal age-dependent mechanism involving reduced recombination. **Human Molecular Genetics**, v. 3, n. 9, p. 1529-1535, 1994.

SHI, J., WOLF, S. E., BURKE, J. M., PRESTING, G. G., ROSS-IBARRA, J., & DAWE, R. K. Widespread gene conversion in centromere cores. **PLoS biology**, v. 8, n. 3, p. e1000327, 2010.

SMITH BJ. boa: an R package for MCMC output convergence assessment and posterior inference. **J Stat Softw**. 21:1–37;2007.

SOUZA, G. L. I. Micro-RNAs espermáticos relacionados à fertilização e ao desenvolvimento embrionário inicial in vitro de bovinos (Bos taurus). 2019.

SUTOVSKY, P. Sperm–oocyte interactions and their implications for bull fertility, with emphasis on the ubiquitin–proteasome system. **Animal**, v. 12, n. s1, p. s121-s132, 2018.

TOLEDO, F., & WAHL, G. M. MDM2 and MDM4: p53 regulators as targets in anticancer therapy. **The international journal of biochemistry & cell biology**, v. 39, n. 7-8, p. 1476-1482, 2007.

TORTEREAU F., SERVIN B., FRANTZ L., MEGENS H. J., MILAN D., ROHRER G., WIEDMANN R., BEEVER J., ARCHIBALD A. L. & SCHOOK L. B. A high-density recombination map of the pig reveals a correlation between sex specific recombination and GC content. **BMC Genomics**, 13:586, 2012.

VANRADEN, Paul M. Efficient methods to compute genomic predictions. **Journal of dairy science**, v. 91, n. 11, p. 4414-4423, 2008.

VERNEKAR, D. V., REGINATO, G., ADAM, C., RANJHA, L., DINGLI, F., MARSOLIER, M. C., & BORDE, V. The Pif1 helicase is actively inhibited during meiotic recombination which restrains gene conversion tract length. **Nucleic Acids Research**, v. 49, n. 8, p. 4522-4533, 2021.

VOUSDEN, K. H., & LU, X. Live or let die: the cell's response to p53. **Nature Reviews Cancer**, v. 2, n. 8, p. 594-604, 2002.

WAGSTAFF, K. M., & JANS, D. A. Importins and beyond: non-conventional nuclear transport mechanisms. **Traffic**, v. 10, n. 9, p. 1188-1198, 2009.

WANG, J., FAN, H. C., BEHR, B., & QUAKE, S. R. Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. **Cell**, v. 150, n. 2, p. 402-412, 2012.

WANG, S., SHANG, Y., LIU, Y., ZHAI, B., YANG, X., & ZHANG, L. Crossover patterns under meiotic chromosome program. **Asian Journal of Andrology**, v. 23, n. 6, p. 562-571, 2021.

WANG, M., XU, Y., ZHANG, Y., CHEN, Y., CHANG, G., AN, G., YANG, X., ZHENG, C., ZHAO, J., LIU, Z., WANG, D., MIAO, K., RAO, S., DAI, M., WANG, D., & ZHAO, X. Y. Deciphering the autophagy regulatory network via single-cell transcriptome analysis reveals a requirement for autophagy homeostasis in spermatogenesis. **Theranostics**, v. 11, n. 10, p. 5010, 2021.

WENG, Z. Q., SAATCHI, M., SCHNABEL, R. D., TAYLOR, J. F., & GARRICK, D. J. Recombination locations and rates in beef cattle assessed from parent-offspring pairs. **Genetics Selection Evolution**, v. 46, n. 1, p. 1-14, 2014.

WENG, Z., WOLC, A., SU, H., FERNANDO, R. L., DEKKERS, J. C., ARANGO, J., & GARRICK, D. J. Identification of recombination hotspots and quantitative trait loci for recombination rate in layer chickens. **Journal of animal science and biotechnology**, v. 10, n. 1, p. 1-13, 2019.

WILSON, M. A., KWON, Y., XU, Y., CHUNG, W. H., CHI, P., NIU, H., MAYLE, R., CHEN, X., MALKOVA, A., SUNG, P., & IRA, G. Pif1 helicase and Polδ promote recombination-coupled DNA synthesis via bubble migration. **Nature**, v. 502, n. 7471, p. 393-396, 2013.

WU, Q., FUKUDA, K., KATO, Y., ZHOU, Z., DENG, C. X., & SAGA, Y. Sexual fate change of XX germ cells caused by the deletion of SMAD4 and STRA8

independent of somatic sex reprogramming. **PLoS Biology**, v. 14, n. 9, p. e1002553, 2016.

WU, Q., FUKUDA, K., WEINSTEIN, M., GRAFF, J. M., & SAGA, Y. SMAD2 and p38 signaling pathways act in concert to determine XY primordial germ cell fate in mice. **Development**, v. 142, n. 3, p. 575-586, 2015.

WU, Q., KANATA, K., SABA, R., DENG, C. X., HAMADA, H., & SAGA, Y. Nodal/activin signaling promotes male germ cell fate and suppresses female programming in somatic cells. **Development**, v. 140, n. 2, p. 291-300, 2013.

YANG, J., LEE, S. H., GODDARD, M. E., & VISSCHER, P. M. GCTA: a tool for genome-wide complex trait analysis. **The American Journal of Human Genetics**, v. 88, n. 1, p. 76-82, 2011.

YANG, S., LIU, Z., WU, S., ZOU, L., CAO, Y., XU, H., HUANG, J., TIAN, Q., WU, F., LI, P., PENG, S., & SHUAI, C. Meiosis resumption in human primordial germ cells from induced pluripotent stem cells by in vitro activation and reconstruction of ovarian nests. **Stem Cell Research & Therapy**, v. 13, n. 1, p. 339, 2022.

YAO, X., CHEN, X., COTTONHAM, C., & XU, L. Preferential utilization of Imp7/8 in nuclear import of Smads. **Journal of Biological Chemistry**, v. 283, n. 33, p. 22867-22874, 2008.

YIN, J., NI, B., TIAN, Z. QIANG, YANG, F., LIAO, W. GONG, & GAO, Y. QI. Regulatory effects of autophagy on spermatogenesis. **Biology of reproduction**, v. 96, n. 3, p. 525-530, 2017.

ZHANG, X. J., HE, J. P., WEN, X. X., & ZHAO, L. Immunolocalisation and expression of Smad2 and Smad4 proteins in dog testis during postnatal development. **Andrologia**, v. 43, n. 4, p. 254-260, 2011.

ZHANG, J., KADRI, N. K., MULLAART, E., SPELMAN, R., FRITZ, S., BOICHARD, D., & DRUET, T. Genetic architecture of individual variation in

recombination rate on the X chromosome in cattle. **Heredity**, v. 125, n. 5, p. 304-316, 2020.

ZHOU, Y., SHEN, B., JIANG, J., PADHI, A., PARK, K. E., OSWALT, A. & MA, L. Construction of PRDM9 allele-specific recombination maps in cattle using large-scale pedigree analysis and genome-wide single sperm genomics. **DNA Research**, v. 25, n. 2, p. 183-194, 2018.

CHAPTER 3 - Variance of gametic diversity estimation for economically important traits of Nellore cattle

ABSTRACT - The aim of this study was to estimate the variance of gametic diversity of Nellore animals for economically important traits using recombination rates between SNP markers, haplotypic phasing information and SNP marker effects and, finally, to compare the response to selection for the traits considering or not the variance of gametic diversity. The reproductive traits considered were scrotal circumference (SC), age at first calving (AFC) and heifer pregnancy (HP), heifer rebreeding (HR) and stayability (STAY), and the growth traits were weaning weight (WW) and yearling weight (YW). Phenotypic records measured in more than 1,100,000 Nellore animals from commercial breeding programs were used in our study. A total of 62,022 genotyped animals imputed from low and medium density panels to a high-density panel (409,029 SNP markers after quality control) were used. The genomic positions of the markers were based on the ARS-UCD1.2 *Bos taurus* genome assembly. Recombination rates were measured previously by an indirect method using haplotypic phasing and pedigree information. A model assuming homogeneity of variance of marker effects, ssGBLUP, was used for all traits. The gamevar.f90 software was used to estimate the variance of gametic diversity and the coefficient of relative variation. Gametic variance was identified across all traits, with BTA1 and BTA8 playing central roles in transmitting genetic variability for growth and reproductive traits, respectively. BTA1 had the greatest impact on the gametic covariance of YW, WWd, and WWm, whereas BTA8 was associated with AFC, HP, and HR, reinforcing their importance in genetic regulation. The inclusion of gametic variance in selection suggested additional genetic gains, with the largest increase observed for HP (+8.41%) and STAY (+8.43%) under REPD1.5. Growth traits showed more modest improvements, ranging from 2.37% to 5.38%. The number of progeny required to capture 90% of the gametic variance varied between 51 (SC and YW) and 125 (HP) in the most influent population sires, highlighting the potential benefits of relative selection to gametic variance with reproductive biotechnologies. These findings highlight the importance of incorporating gametic diversity into genetic evaluations and reinforce its value as a complementary criterion to traditional selection methods, enhancing both the accuracy and long-term efficiency of Nellore breeding programs.

Keywords: beef cattle, coefficient of relative variation, haplotypes, Mendelian sampling, recombination rate

3.1 INTRODUCTION

Genetic improvement programs aim to increase the frequency of favorable alleles in a population, thereby promoting progress in economically important traits. Traditionally, the selection of superior individuals has been based on estimated breeding values (EBVs), which are derived from phenotypic records and pedigree information. The emergence of large-scale genotyping technologies represented a major advancement in this field, enabling the integration of molecular marker data into genetic evaluations. This innovation led to the development of genomic estimated breeding values (GEBVs), which combine phenotypic, pedigree, and genomic information. As a result, GEBVs provide greater accuracy in selection decisions and contribute to shorter generation intervals, enhancing the overall rate of genetic progress (MEUWISSEN et al., 2001; SCHAEFFER, 2006).

Despite these advancements, both EBVs and GEBVs summarize the average additive genetic effects, overlooking the variability introduced by meiotic processes such as recombination and Mendelian sampling (MÜLLER et al., 2018; BIJMA et al., 2020; SANTOS et al., 2019). These methodologies limit their capacity to predict the genetic diversity transmitted to offspring, which, in turn, is an important factor for sustaining long-term genetic progress and diversity.

Gametic diversity, arising from these meiotic processes, is a key component for optimizing breeding strategies. It accounts for the diversity of gametes produced by an individual, which can influence the variation of progeny performance. Segelke et al. (2014) and Santos et al. (2019), highlight the importance of considering gametic variability, as even individuals with high GEBVs can produce offspring with varying genetic potential due to differences in the genetic composition of their gametes. To address this, the concept of gametic diversity variance has been incorporated into selection models, providing a more comprehensive framework for mating decisions (SANTOS et al., 2019). By balancing genetic gain with the preservation of genetic diversity, this approach enhances the sustainability of breeding programs (SANTOS et al., 2019; VANRADEN, 2020; BIJMA et al., 2020).

Several studies have explored ways to integrate gametic diversity into selection programs. For example, Tabet (2022) demonstrated through simulations that including gametic variance in selection indices can maintain higher genetic variability over multiple generations, reduce inbreeding rates, and improve long-term genetic gain. Similarly, Niehoff et al. (2024a, 2024b) proposed methodologies that incorporate Mendelian sampling variance into selection decisions, leading to greater retention of genetic variability and higher genetic gain over multiple generations. These studies, along with the analysis by Bijma et al. (2020), reinforce that integrating gametic diversity variance is essential for the sustainability and long-term success of genetic improvement programs.

While many of these studies have focused on dairy cattle or other species, there are no works that investigated the application of gametic variance in the Nellore breed, which is important for beef production in tropical environments. Considering that Brazil is the world's largest exporter of beef, and that the Nellore breed represents a portion of the national herd (ABIEC, 2023), applying these innovative approaches could be essential to optimizing the breed's improvement programs, promoting both genetic gain and the preservation of genetic diversity (LOZADA-SOTO et al., 2021). The aim of this study was to estimate the variance of gametic diversity of genotyped Nellore animals for economically important traits using recombination rates between SNP markers, haplotype phasing information and SNP marker effects and, finally, to compare the response to selection for those traits considering the variance of gametic diversity and the traditional GEBV.

3.2 MATERIAL AND METHODS

The Nellore database belonged to three commercial breeding programs (DeltaGen, Paint and Cia de Melhoramento), which are part of Alliance Nellore database (www.gensys.com.br). The animals, born between 1984 and 2019, were from 276 commercial herds widely distributed in the Midwest, Southeast, and Northeast of Brazil with high connectedness by the common sires intensively used through artificial insemination (AI), with more than 50% of the calves born from AI. The animal procedures realized in this research agreed with Animal Care

of the São Paulo State University (UNESP), School of Agricultural and Veterinary Science Ethical Committee (protocol number 18.340/16).

3.2.1 Genotypic data information

Genotype inconsistences between parents and progeny were adjusted using conflict.f90 software, which corrects for mendelian errors and fills missing SNP using parental genotypes (VANRADEN et al., 2015). A total of 62,022 Nellore animals, 24,042 females and 37,980 males genotyped with different Bead chip assay densities according shown in Table 1. The males and females genotyped with the lower density panel were imputed to the HD panel using FImpute v3 (Sargolzaei et al., 2014) with expected accuracy higher than 0.97 (CARVALHEIRO et al., 2014).

The QCF90 software (MASUDA et al., 2019) was used for the quality control of SNP markers and samples. SNPs located in the autosomes with a GeneCall score higher than 0.80 were considered in the analyses. In addition, SNPs located at the same genomic position, with MAF (Minor allele frequency) ≤ 0.05 and Call rate ≤ 0.98 were removed considering the ARS-UCD1.2 *Bos taurus* genome assembly. A threshold of 0.16 was used for Hardy-Weinberg equilibrium, tested by the maximum deviation of heterozygote frequency from expected, and SNPs that exceeded this value were excluded. Moreover, samples with a call rate lower than 0.90 were also removed from the analyses After editing 62,022 genotype animal and 409,029 SNPs remained in genomic data set to carry out further analyses.

Table 1 – Number of animals and SNP markers before quality control of Nellore animals genotyped in the study

Number of males	Number of females	SNP number	Chip	Manufacturer
721	307	17,794	Zoetis ZL4 (Kalamazoo, MI, USA)	Zoetis
145	-	19,720	GeneSeek® Genomic Profiler 20K – Indicine (Lincoln, NE, USA)	Neogen

49	-	26,151	GGP Bovine LDv3 (Lincoln, NE, USA)	Neogen
6,617	4,210	27,553	Zchip (Araçatuba, SP, BR)	Deoxi
11,703	2,878	29,842	Zoetis Custom SNP ZL5 (Kalamazoo, MI, USA)	Zoetis
1	54	30,108	GGP Bovine LDv3 (Lincoln, NE, USA)	Neogen
10,526	9,334	35,339	GGP Indicus 35K (Lincoln, NE, USA)	Neogen
870	429	54,609	Illumina BovineSNP50 (San Diego, CA, USA)	Illumina
1,569	2,198	54,791	GGP Indicus 50K (Lincoln, NE, USA)	Neogen
3,556	1,372	74,153	GeneSeek® Genomic Profiler 75K – Indicine (Lincoln, NE, USA)	Neogen
2,223	3,260	777,962	Illumina BovineHD BeadChip (San Diego, CA, USA)	Illumina

3.2.2 Phenotypic information

Phenotypic information measured on approximately 1,100,000 Nellore animals was provided by commercial breeding programs from the Alliance Nellore database (www.gensys.com.br). Two sets of traits, growth and reproductive traits, were analyzed (Table 2). Reproductive traits are scrotal circumference (SC), age at first calving (AFC) and heifer pregnancy (HP), heifer rebreeding (HR) and stayability (STAY). The growth traits consisted of weaning weight (WW) and yearling weight (YW).

The animals were reared in grazing systems with varying nutritional levels. In some herds, the animals received protein and mineral supplementation, especially during the dry season, while in others, only urea supplementation was provided. Generally, the heifers were exposed to two breeding seasons: an early breeding season, typically occurring in the first trimester (January to March) for 60 days, during which heifers were exposed to their first mating at around 16 months of age (regardless of body weight and body condition score). Heifers that did not conceive during the early breeding season were given a second

opportunity during the regular breeding season (November to January), approximately 8 months later, at around 24 months of age. This early breeding test, initiated in the 1990s (FERNANDES et al., 2022), is conducted to identify precocious heifers in out-of-season breeding and generate data for use in genetic evaluations. Additionally, some breeders began exposing females to reproduction around 14 months of age during a 90-day breeding season, aiming to intensify the selection process for sexual precocity. Reproduction is carried out through artificial insemination or natural mating. When a fixed-time artificial insemination protocol was used, the entire contemporary group followed the same protocol. Pregnancy diagnosis in heifers is performed approximately 60 days after the end of the breeding season; females that did not conceive in either breeding season were culled.

The AFC was computed as the difference between the date of first calving and heifer date of birth, in days. HP was defined as attributing a value of 1 (success) to heifers calving until 30 months of age and a value of 0 (failure) otherwise. HR, which was determined attributing a value of 1 (success) or 0 (failure) to females that presented or not a second calf, respectively. STAY was determined by attributing a value of 1 (success) to cows with at least three calving's up 76 months and, otherwise, a value of 0 (failure). SC was measured at 365 days of age, using a specific metal tape at the widest point of the scrotum, in cm. Animals WW (kg) were measured at weaning with 221 days of age, on average, and YW (kg) was measured with 15 months of age (450 d), on average.

Contemporary groups (CG) for AFC, HP, HR, STAY, SC and YW considered animals from the same year and season of birth, herd (at birth, weaning, and yearling), and management group at birth, weaning, and yearling. For WW, considered animals from the same year and season of birth, herd (at birth and weaning), and management group at birth and weaning. The season of birth was divided into two from August to January and from February to July. The CG with less than five animals in which all scores for HP, HR and STAY were the same, i.e., groups without variability, were eliminated from the analysis to avoid extreme category problems (MISZTAL et al., 1989). For continuous traits, observations outside the interval of 3.5 standard deviations below and above the CG mean were excluded.

3.2.3 Statistical models and Analysis

The (co)variance components were obtained by Gibbs sampling, using GIBBSF90 and THRGIBBSF90 programs (MISZTAL et al., 2002). A threshold animal model was used to analyze HP, HR and STAY, whereas a linear animal model was used for the other traits (AFC, SC, WW and YW). The posterior estimates were obtained using the POSTGIBBSF90 program (MISZTAL et al., 2002). Analysis consisted of a single chain with length of 1,000,000 cycles, in which the first 200,000 cycles were discarded (burn-in), considering a sampling thinning interval of 20 cycles. Thus, 40,000 samples were effectively used for final inferences (Table 2).

The models used for variance components estimation and breeding value prediction included the systematic effects of CG (calf sex and age class of the cow were included as fixed effects only for AFC) and covariates (linear and quadratic effects for age of the animal at recording and age of the dam at calving were included for WW, YW, and SC). For HP, HR, and STAY, the female age at recording was included as a covariate (linear and quadratic effects). Additionally, direct additive genetic, maternal genetic, maternal permanent environmental, and residual random effects were included for WW. For SC, AFC, HP, HR, STAY, and YW, only the direct additive genetic and residual effects were considered as random.

The following liability function (SORENSEN & GIANOLA, 2002) was used to analyze HP, HR and STAY:

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

where: y_i is the binary observation (categories 0 or 1), l_i is the underlying liability of observation i, t_i is the threshold that defines the response category for the trait, and n_i is the total amount of data for studied trait. Observable data are dependent on the underlying variable, which is limited by one unobservable threshold (t_i). For example, if the unobservable value is $<t_i$ the response variable is 0, otherwise the value is 1.

Table 2 - Description of the dataset used in the analyses and heritabilities for reproductive and growth traits in Nellore cattle

Traits	Item				
	Nº of animals with record	Nº of animals in the pedigree	Nº of contemporary groups	Direct heritability	Maternal heritability
AFC, days	357,775	567,770	19,303	0.31	-
HP, 0 or 1	159,872	326,837	8,253	0.15	-
HR, 0 or 1	300,831	496,224	15,667	0.12	-
STAY, 0 or 1	190,243	327,235	10,839	0.21	-
SC, cm	531,579	973,541	16,336	0.37	-
WW, kg	1,101,833	1,553,073	29,485	0.16	0.16
YW, kg	1,102,880	1,553,988	33,607	0.37	-

Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight (WW); Yearling weight (YW)

3.2.4 Estimation of SNP marker effects

To estimate marker effects, we used a genomic model that assumes homogeneity of variance of marker effects through a Single-step approach (ssGBLUP) (MISZTAL et al., 2009). In the ssGBLUP approach, the inverse of the relationship matrix ($\mathbf{A}^{-1}$) based on pedigree is combined with a genomic relationship matrix into a genomic-pedigree relationship $\mathbf{H}^{-1}$ (AGUILAR et al., 2010):

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

where, $\mathbf{H}^{-1}$ is the inverse of the modified relationship matrix; $\mathbf{A}_{22}^{-1}$ is the inverse of the pedigree relationship matrix for genotyped animals; and $\mathbf{G}^{-1}$ is the inverse of the genomic relationship matrix (VANRADEN, 2008), which is described as:

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

where **W = (M – P)**, in which **M** is the SNP incidence matrix, with m columns representing the number of markers and n lines representing the number of genotyped animals. Each element in **M** was set to 0, 1, or 2, for genotypes aa, Aa, and AA, respectively. **P** is the matrix containing the allele frequencies expressed in $2p_i$, where p_i is the frequency of the second allele at locus i.

The SNP effects were obtained based on the genomic values (GEBV) of the genotyped animals, using the POSTGGSF90 software (AGUILAR et al., 2014). The equation for computing the effect of SNP can be described as:

$$\hat{u} = Z'G^{*-1}\hat{a}_g = Z'\left[ZZ'\right]^{-1}\hat{a}_g$$

where $\hat{u}$ is the effect vector of SNPs effect; Z' is the transposed matrix that relates the genotypes of each locus; G^{*-1} is the inverse matrix of weighted genomic relationships; $\hat{a}_g$ is the vector with the predicted additive genetic values of the genotyped animals (GEBVs).

3.2.5 Estimation of the Gametic Diversity Variance

The gamevar.f90 software (SANTOS et al., 2020) was used to estimate variance of gametic diversity (σ^2_{gamete}). Considering the recombination information between SNPs required for the analysis, a recombination map for Nellore animals, previously computed, was used. In addition, the data was phased by the findhap.f90 v4 software. The σ^2_{gamete} represents the variability of all possible gametic values generated by the permutation and recombination of each parental chromosome, and only the heterozygous loci of an individual contribute to σ^2_{gamete} (SANTOS et al., 2019). This can be represented in matrix format as follows:

$$\sigma^2_{gamete} = [\alpha_1 \cdots \alpha_N] M [\alpha_1 \cdots \alpha_N]',$$

where $\alpha_j (j = 1, \ldots, N)$ are the allele substitution effects, and **M** is the (co)variance matrix of the Mendelian transmission probabilities for the *N* heterozygous *loci*:

$$\boldsymbol{M} = \begin{bmatrix} 0{,}25 & \cdots & al_{1,N}\left(-\frac{rate_{1,N}}{2} + 0.25\right) \\ \vdots & \ddots & \vdots \\ al_{N,1}\left(-\frac{rate_{N,1}}{2} + 0.25\right) & \cdots & 0{,}25 \end{bmatrix},$$

where $al_{jk}$is a phase indicator for loci j and k, with value 1 when both loci have the reference allele on the same chromosome and −1 otherwise; $rate_{i,j}$ is the genetic distance between the 2 *loci* (in centimorgans). All *loci* with recombination > 0.50 on the same chromosome, or on different chromosomes, are assumed to be independent and thus have zero values for the corresponding elements of ***M*** (SANTOS et al. 2019).

The gamevar.f90 software (SANTOS et al., 2020) was also used to estimate the coefficient of relative variation (CRV) that measures the variability in the percentage of additive genetic values transmitted from an individual to its gametes (GEBV/2), which is useful in breeding and progeny testing programs to estimate the optimal number of progenies needed to find the expected gametic variability (SANTOS et al., 2019). The CRV of an individual *i* is defined as follows:

$$CRV_i = \frac{\sigma_{gamete}}{\sqrt{0.5 \sum_h^{NHom} \alpha_h^2 + \sigma^2_{\text{gamete}}}}$$

Where $\sum_h^{NHom} \alpha_h^2$ is sum of squared effects of the homozygous loci from an individual.

3.2.6 Application of Gametic Variance in Selection Programs

Focusing on the future genetic progress, we used a strategy proposed by Santos et al. (2019), that proposed a selection criterion including the gametic

variance relative to the selection intensity applied in the de next generation (i_f) for an individual i (unknowing mating),

$$REPD_i = EPD_i + \sigma_{gamete\_i} \times i_f,$$

where *REPD$_i$* (relative EPD - (Expected Progeny Difference) is the average of the genetic values relative to the group of progeny that will be selected in the future. This index was used for all traits except for AFC, where a modification in the gametic term adapted for subtracting it from the EPD, since lower values are desirable in this trait. The REPD is a measure with biological interpretation, whose value represents the expected difference of the selected gametes, in relation to the genetic base of the population, when a certain selection intensity is applied to all gametes of an individual. To verify the impact of using REPD instead of traditional EPD as selection criterion, we estimated the rank correlation between REPD and EPD; calculated the percentage of coincidence for different reposition rate and obtained the expected genetic gain by using REPD and traditional EPD.

3.3 RESULTS AND DISCUSSION

This study analyzed seven economically traits in Nellore cattle, encompassing reproductive traits (scrotal circumference - SC, age at first calving - AFC, heifer pregnancy - HP, heifer rebreeding - HR, and stayability - STAY) as well as growth traits (weaning weight - WW and yearling weight - YW). Among these, HP, HR, and STAY were treated as categorical traits, and their analyses employed the liability scale, enabling a robust interpretation of gametic diversity variance. Additionally, one trait was analyzed gametically for maternal effects, introducing another innovative component to the study.

As reported by Santos et al. (2019), no significant differences were identified between sexes regarding gametic diversity variance of dairy traits. This finding indicates that genetic variability among gametes is not influenced by the sex of individuals, suggesting the feasibility of the use of the male recombination

map to estimate gametic diversity across the entire population without the need for sex-based stratification. Utilizing this map allowed for precise capture of variability among gametes, directly reflecting the impact of recombination, including the hotspot regions, and the Mendelian sampling on the genetic diversity observed in offspring. As a pioneering analysis of gametic diversity variance in Nellore cattle, this study highlights the importance of assessing all traits of interest, including categorical ones, and monitoring genetic diversity at every stage of genetic planning.

3.3.1 Obtaining the variance of gametic diversity and related parameters

Variance of gametic diversity and CRV in the whole Nellore population

Similarly to Santos et al., 2019, firstly we calculated the gametic variance separately for each chromosome, since they are independent genome structures. The average, standard deviation, and amplitude of the estimated gametic variance were well distributed among chromosomes (Figure 1; Table S1). For STAY, SC and non-reproductive traits (YW, WWd and WWm), chromosome 1 explained greater variation of the distribution of the gametic variance, while for reproductive traits (AFC, HP, HR) chromosome 8 slightly standout.

Genome-wide association studies (GWAS) have identified chromosome 8 as a key region associated with reproductive traits in cattle. In Bos indicus breeds, Tahir et al. (2021) reported a 0.5 Mb region containing the genes *SLC44A1, FSD1L, FKTN, TAL2,* and *TMEM38B*. The homologous region in humans, located on chromosome 9, has been linked to female sexual maturation age (Perry et al., 2009). Additionally, *SLC44A1* and *TMEM38B* have been associated with age at first sex and pubertal growth (MILLS et al., 2020; COUSMINER et al., 2014). Further studies have identified quantitative trait loci (QTL) and functional genes on chromosome 8 related to reproductive traits. Michenet et al. (2016) found a QTL with strong evidence in this region, where the *PALM2* gene, involved in cell formation and plasma membrane dynamics (KUTZLEB et al., 1998), was located near the peak SNP. Another gene, *RGP1*, mapped within the same region, plays a role in converting guanosine diphosphate (GDP) into guanosine triphosphate (GTP), a critical process for protein activation in cellular metabolism

and milk production (TURNER et al., 1992). Chromosome 8 has also been associated with key reproductive processes in dairy cattle. Fritz et al. (2013) identified deleterious haplotypes in the *GART, SHBG,* and *SLC37A2* genes, linked to embryonic loss in Holstein cattle. Similarly, Daetwyler et al. (2014) mapped variants on this chromosome associated with both monogenic and complex traits, reinforcing its importance for fertility and other productive traits.

In our study on Nellore cattle, we observed increased variance in gametic diversity on chromosome 8 for female reproductive traits, including AFC, HP, HR and STAY (Figure 1). This suggests that chromosome 8 harbors loci influencing fertility, aligning with findings from both beef and dairy breeds. Despite differences in selection criteria and breeding objectives, the recurring association of this chromosome with reproductive traits highlights the need for further investigation into its role in reproductive efficiency in Nellore cattle. The presence of multiple genomic regions and candidate genes linked to fertility on chromosome 8 underscores its importance in cattle breeding. Identifying functional variants and understanding their effects could help improve pregnancy rates and reduce AFC, both crucial factors for enhancing productivity in tropical beef cattle production systems.

The influence of BTA1 (*Bos taurus* autosome) on the variance of gametic diversity for SC (scrotal circumference) and STAY (stayability) (Figure 1) may be associated with key genes involved in testicular function and hormonal regulation. Teixeira et al. (2017) identified *IGSF11* in BTA1, a gene from the immunoglobulin superfamily that is highly expressed in testicular tissue. Suzu et al. (2002) demonstrated that *IGSF11* plays a role in neuroendocrine regulation, influencing hormonal pathways such as LH, FSH, and prolactin, which are essential for both male fertility and female reproductive longevity. Additionally, Neves et al. (2019) identified *IGSF11* as a candidate gene associated with testicular hypoplasia (TH) in Nellore cattle, reinforcing its potential role in testicular development, sperm production, and scrotal circumference.

Another gene identified in the same BTA1 region is RERG (Ras-like, Estrogen-Regulated, Growth Inhibitor), which belongs to the Ras superfamily of GTPase proteins and acts as a growth regulator influenced by estrogen. Neves et al. (2019) demonstrated that *RERG* is expressed in endocrine tissues, including the adrenal glands, thyroid, and testis, suggesting its potential influence

on steroid hormone biosynthesis. Since the adrenal glands play a fundamental role in producing hormones essential for reproductive regulation, the presence of *RERG* in BTA1 suggests a connection between hormonal regulation, testicular development, and reproductive efficiency.

The coexistence of *IGSF11* and *RERG* in BTA1, along with their functional implications in hormonal regulation and testicular growth, reinforces the importance of this chromosome in male fertility (SC) and female reproductive longevity (STAY). These findings suggest that genomic selection based on gametic diversity variance in BTA1 could be a promising approach to optimizing sexual precocity and reproductive persistence in Nellore cattle.

The strong influence of BTA1 on growth traits YW, WWd, and WWm (Figure 1) may be associated with genes previously identified in this region. Morsci et al. (2006) mapped quantitative trait loci (QTLs) on BTA1 linked to weaning weight and carcass yield in Angus cattle, highlighting *ADIPOQ* (Adiponectin) and *SST* (Somatostatin) as candidate genes. *ADIPOQ* regulates lipid metabolism and fat deposition, influencing muscle growth and feed efficiency, while *SST* modulates the release of growth hormone (GH) and insulin-like growth factor 1 (IGF-1), both essential for bone and muscle development.

Another key gene located on BTA1 is *POU1F1* (Pituitary-specific positive transcription factor 1), positioned in the proximal region of the chromosome, as described by Cardoso et al. (2014). This gene regulates the development of the anterior pituitary gland, influencing the production of GH, prolactin, and thyroid-stimulating hormone (TSH). Mutations in *POU1F1* have been associated with anterior pituitary hypoplasia, GH deficiency, and dwarfism in mammals (LI et al., 1990; PFÄFFLE et al., 1992), reinforcing its role in growth regulation.

The presence of these genes on BTA1 indicates that this chromosome harbors important genetic targets for selecting Nellore cattle with enhanced growth performance. Utilizing gametic diversity variance to identify individuals with greater genetic potential for YW, WWd, and WWm could be an effective strategy to optimize breeding programs, promoting predictable growth adapted to tropical conditions.

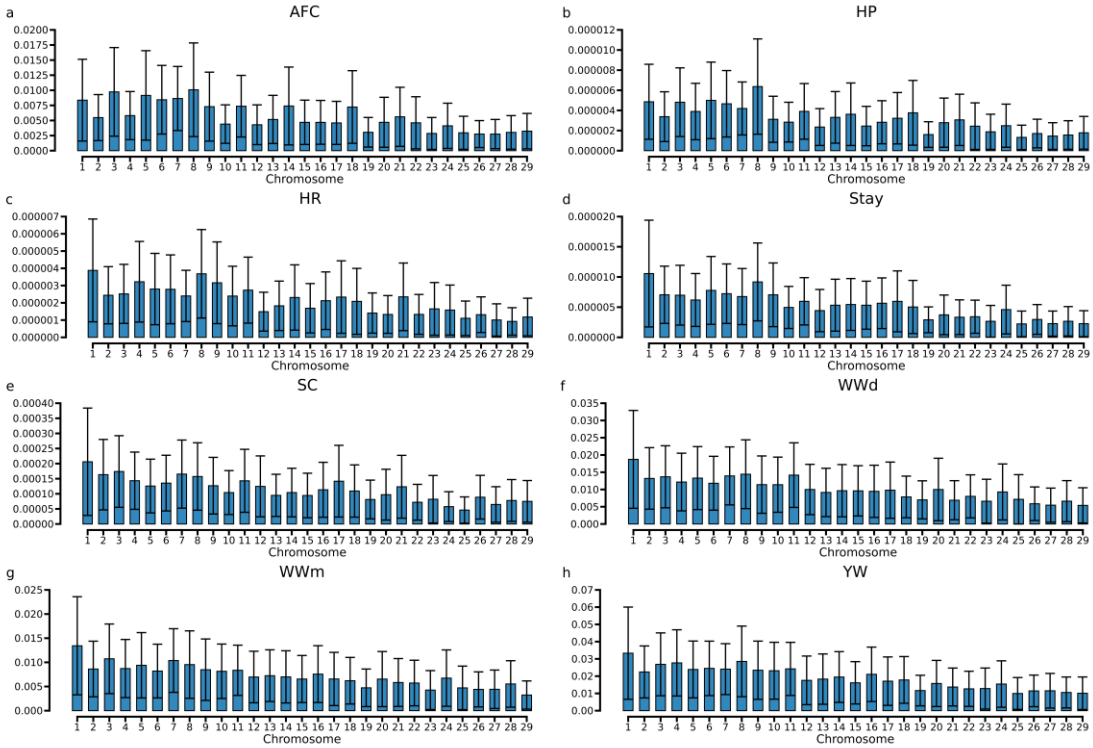
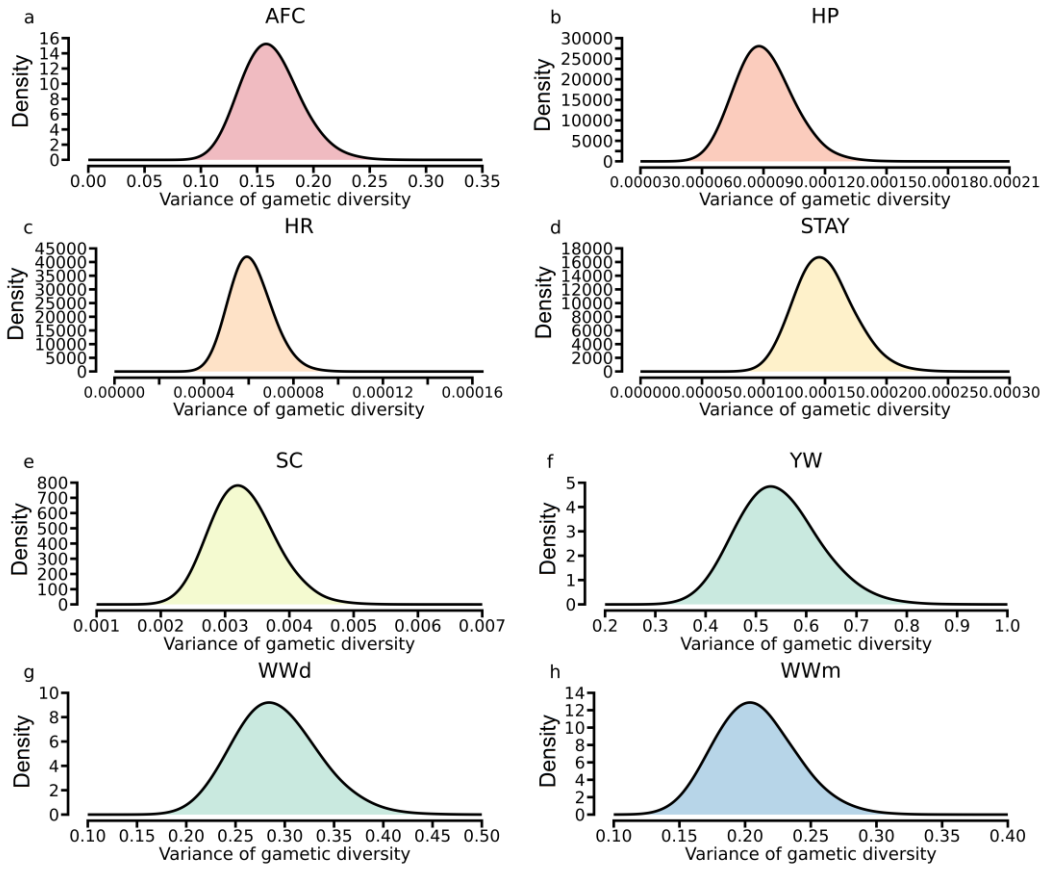


Figure 1 - Bar plots (distribution) of variance of gametic diversity by chromosome for different traits. Bars indicate averages and whiskers represent standard deviations; Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)

We estimated the gametic variance and other related statistics of AFC, HP, HR, STAY, SC, BWY, WWd and WWm for a Brazilian Nellore population. There is a considerable amount of variation in gametic diversity across the animals (Figure 2), which indicates the potential of applying gametic selection to the beef cattle population. In Table S1 is presented descriptive parameters, which can be useful for future simulations studies. The results showed a distribution close to the typical Gaussian curve. Santos et al., 2019 observed similar results for only for milk protein yield in Dairy cattle. But these authors also found atypical curves for Milk Yield, Fat Yield and fat and protein percentage due to the influence of the gene DGTA1 on BTA 14. Thus, the observed distribution in our study is explained by the absence of genes with pronounced mayor effect on the reproduction and beef traits.

Although they visually tend to a normal, through a statistical normality test (Kolmogorov-Smirnov test) it was possible to verify that the tendency towards a Gaussian distribution of the gametic variance of the studied population differs significantly from anormal distribution (p value<0.05). However, for the distribution of CRV, no significant differences from a normal distribution were found (p value>0.05) (Table S2 – Significant values – non-normal; non-significant values - normal).



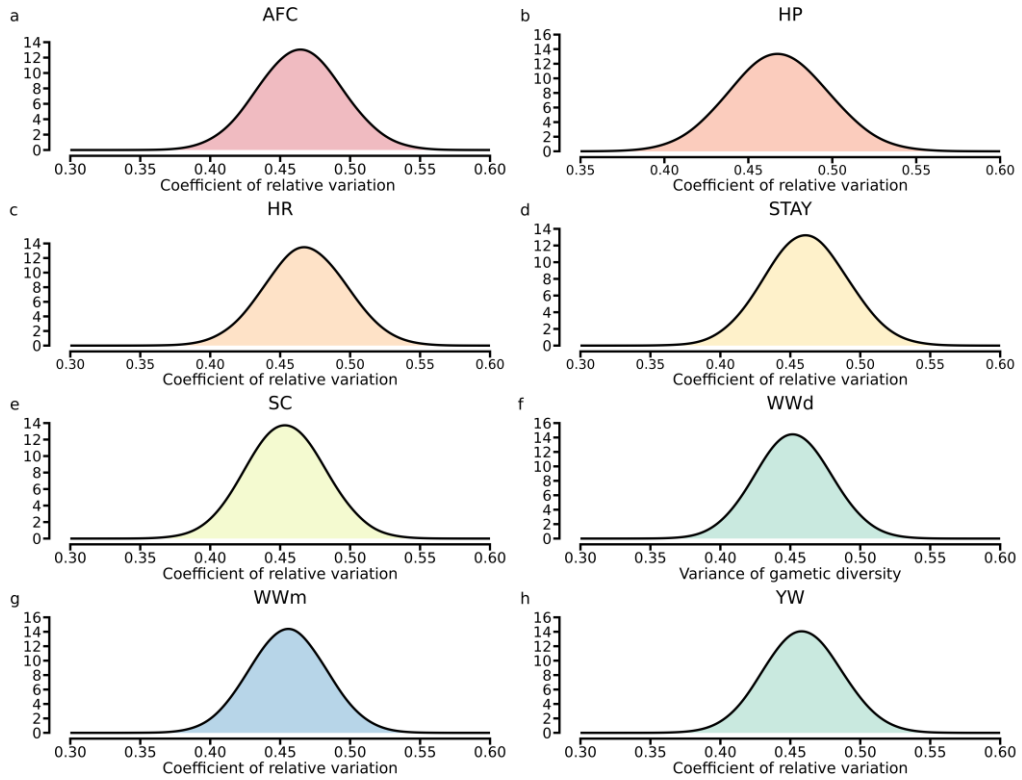


Figure 2 - Kernel density plots of variance of gametic diversity and CRV (coefficient of relative variation) distribution for different traits; Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)

Table 3 presents Pearson correlations between gametic diversity variance (above the diagonal) and breeding values (below the diagonal) for different traits. As expected, the correlations between EBV and gametic variance along the diagonal are close to zero, confirming the findings of Santos et al. (2019) and Segelke et al. (2014). This result aligns with theoretical expectations, as breeding values reflect only expected additive genetic effects, while gametic diversity variance captures additive variation due to Mendelian segregation, which contributes to genetic diversity in offspring.

Santos et al. (2019) and Segelke et al. (2014) reported that Pearson correlations between gametic diversity variance and different productive and reproductive traits vary depending on the genetic structure of the population and the influence of specific genes. Santos et al. (2019) observed moderate to high

Pearson correlations for productive traits in dairy cattle, highlighting that some followed normal distributions, while others showed atypical patterns because of major genes, such as *DGAT1* on chromosome 14. Meanwhile, Segelke et al. (2014) identified positive correlations between the standard deviation of gametic breeding values (SDGBV) for different traits, although these were generally lower than the correlations observed for the mean gametic breeding values (MGBV). These findings indicate that gametic diversity variance may exhibit distinct correlation patterns depending on the trait analyzed and the impact of specific genomic variants.

In the present study, Pearson correlations between gametic diversity variance for different traits showed patterns like those reported by these authors (SEGELKE et al., 2014; SANTOS et al., 2019), with some moderate to high correlations and others close to zero. The highest correlation was observed between age at first calving and heifer pregnancy (0.65), indicating that individuals with higher gametic diversity variance for age at first calving also exhibit higher gametic diversity variance for heifer pregnancy. This result suggests that gametic variability in these traits may be associated with common mechanisms regulating fertility. Additionally, the moderate correlation between weaning weight direct and weaning weight maternal (0.53) suggests that gametic diversity variance for growth traits may be associated with significant maternal effects, as the maternal ability of cows directly impacts calf performance.

The correlations between yearling weight (YW) and both direct and maternal weaning weight also provide relevant insights (Table 3). The moderate correlation between YW and direct weaning weight (0.26) suggests that gametic diversity variance may show some degree of continuity between these growth phases. However, the slight lower correlation between YW and maternal weaning weight (0.19) indicates that maternal effects on weaning weight do not extend as strongly to yearling weight, suggesting that maternal influence on growth may be more pronounced in the early stages of development.

Overall, the presence of Pearson correlations closes to zero between certain trait combinations indicate that gametic diversity variance between them is independent in the population level, meaning there is no significant interdependence in how gametic diversity variance is overall transmitted for these traits. This reinforces the importance of considering gametic diversity variance as

a complementary tool for selection, allowing the evaluation of which traits have greater potential response to selection based on the variation present in gametes.

The results in Table 3 thus reinforce that gametic diversity variance can provide additional information for genetic improvement programs, particularly for reproductive traits. For growth traits, weaker correlations indicate that different mechanisms influence direct and maternal effects, emphasizing the need for multifactorial approaches in Nellore cattle selection, such as the use of relative indices. These findings align with the studies of Santos et al. (2019) and Segelke et al. (2014), reinforcing the importance of considering gametic diversity variance in genomic selection to enhance the genetic variability available in the population.

Table 3 - Pearson correlations between estimates of variance of gametic diversity for different traits (above the diagonal), breeding values (below the diagonal), and correlations between variance of gametic diversity and breeding values for each trait (diagonal)

Traits	AFC	HP	HR	STAY	SC	YW	WWd	WWm
AFC	**0,02**	0,65	0,06	0,12	0,06	0,01	0,02	0,04
HP	-0,81	**-0,01**	0,08	0,09	0,04	0,04	0,04	0,05
HR	-0,23	0,17	**-0,01**	0,18	0,03	0,03	0,03	0,03
Stay	-0,10	0,06	0,38	**-0,03**	0,06	0,00	0,03	0,03
SC	0,03	-0,05	0,01	-0,02	**0,00**	0,08	0,04	0,04
YW	-0,04	0,07	0,03	0,05	0,18	**-0,02**	0,26	0,19
WWd	-0,10	0,06	0,06	0,06	0,12	0,43	**-0,01**	0,53
WWm	-0,11	0,11	0,04	0,02	0,15	0,33	0,72	**-0,01**

Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)

Considering the 1,978 sires (males with genotyped offspring), we estimated the number of progenies required to capture 90% of the gametic variance (or to allow, at most, a 10% variation in the EBV predicted using only progeny data) in future generations (Figure 4). The estimated values varied

across the evaluated traits, ranging from 51 progenies for SC and YW to 125 for HP.

The number of progenies was calculated following the methodology of Santos et al. (2020):

$$n = \frac{(1.96)^2 x (CRV_i)^2}{(0.1)^2}$$

The variation in the estimated number of progenies was very high across traits, ranging from 50 to 125, indicating that the EPDs of some sires would require much less or much more offspring information to stabilize in 90%, considering 95% confidence. Furthermore, this pattern can be explained by the differential heterozygosity of loci affecting specific traits, which directly influences the CRV estimate and variation of EDPs. This finding aligns with the results of Segelke et al. (2014), who reported that Mendelian sampling varies widely across traits and individuals, and that contributes significantly to selection accuracy. However, the correlations between the estimates exhibited high, and mainly moderate to low values (Table 4), indicating that the same sire may require fewer progenies for some traits and more for others. This highlights the prioritization of objective traits and the use of economic weight even in the first step of genetic evaluation designs. The highest correlation was observed between AFC and HP (0.63), and direct weaning weight (WWd) and maternal weaning weight (WWm) (0.53), suggesting that sires requiring a larger number of progenies for one of these traits tend to require more for the other as well. Additionally, moderate correlations between YW and WWd (0.27) and between YW and WWm (0.21) suggest that gametic variance follows a continuous pattern throughout animal development. However, reproductive traits such as HP, HR, and STAY showed low or null correlations with growth traits, indicating that these trait groups have largely independent genetic architectures.

Another important finding was the significantly higher number of progenies required for HP (125) compared to SC (51), reinforcing the idea that fertility traits have a more complex genetic architecture influenced by multiple environmental and genetic factors. Probably, low-heritability traits, such as fertility, rely more on Mendelian segregation and thus require larger sample sizes to accurately estimate gametic variance. Santos et al. (2019) also pointed out that variability in

mendelian among individuals directly affects the required number of additional progenies.

Since different traits require distinct numbers of progenies to capture genetic variability, it is essential that sire selection considers not only the estimated average genetic value but also gametic variance. As proposed by Niehoff et al. (2024b), optimizing sire selection can minimize the need for large progeny populations, ensuring efficient use of genetic resources while maintaining sustainable genetic variability.

Thus, if a general progeny test design or gametic evaluation is desired, certain traits should be prioritized to optimize breeding program efficiency. This highlights the importance of economic selection indices, which can encompass genomic regions of interest, such as Net Merit for dairy cattle (SANTOS et al., 2020). Integrating such indices into selection programs allows for a more efficient strategy, maximizing genetic progress without compromising gametic variance and genetic diversity within the population.

The findings of this study demonstrate that incorporating gametic variance into breeding strategies enables the identification of sires that not only exhibit high genetic values but also transmit useful genetic variability for selection response across generations. Therefore, the results align with recent literature and reinforce the potential of gametic diversity as a key tool for optimizing the efficiency of selection programs in Nellore cattle.

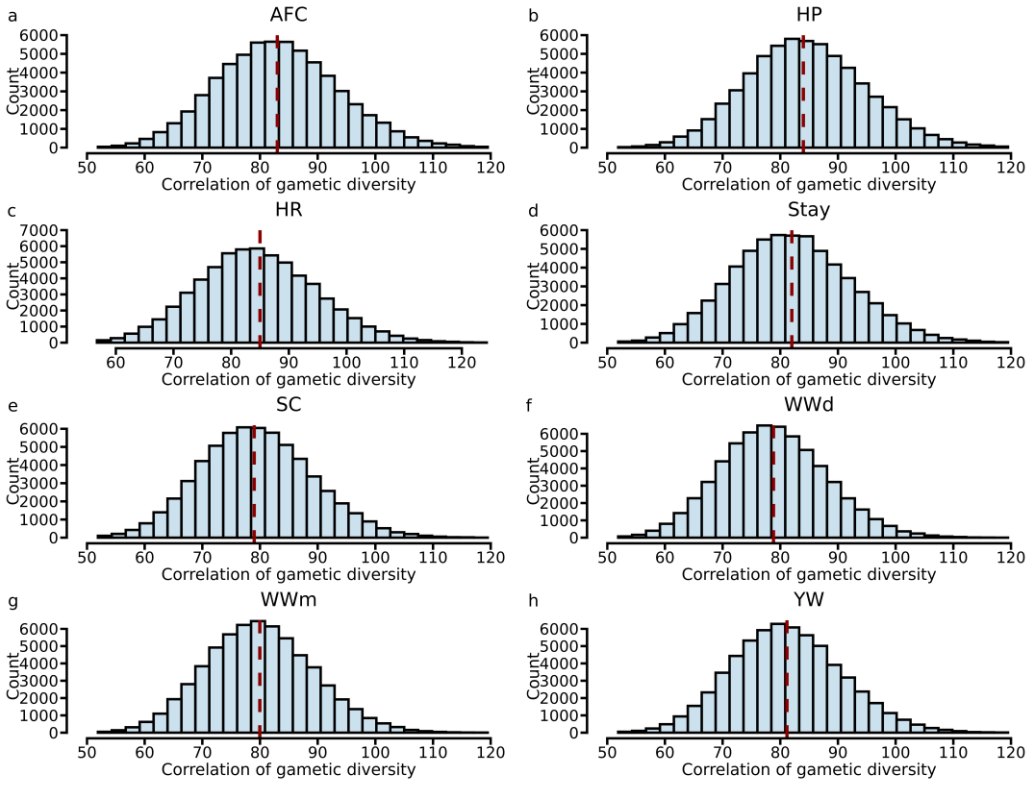


Figure 4 - Histograms of the number of progeny required to realize 90% of gametic variance (or to allow at maximum 10% of change in the EBV predicted using only progeny data) in the future progeny for 1978 sires; Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)

Table 4 - Pearson correlation of the number of progeny (below diagonal) and variance of gametic diversity (above diagonal) between traits considering 1978 sires

	AFC	HP	HR	STAY	SC	YW	WWd	WWm
AFC	-	0,65	0,09	0,13	0,030302	-0,00013	0,012436	0,02173
HP	6,55E-01	-	0,09	0,10	0,020387	0,040914	0,011353	0,031303
HR	1,27E-01	1,26E-01	-	0,16	0,034743	0,026829	-0,00599	0,014158
STAY	1,51E-01	1,56E-01	1,95E-01	-	0,03364	-0,01902	0,025614	0,016262
SC	6,07E-02	5,13E-02	7,04E-02	6,37E-02	-	0,083269	0,079197	0,077607
YW	2,83E-02	7,02E-02	5,81E-02	1,59E-02	0,110338	-	0,259413	0,197085
WWd	4,42E-02	4,46E-02	2,83E-02	6,33E-02	0,108825	0,275702	-	0,522856
WWm	0,05	0,06	0,04	0,05	0,10209	0,216046	0,533973	-

Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)

3.3.2 Correlation of Gametic Variance in the whole population

The analysis of covariance (by chromosome of individuals) and correlation (by individuals) of gametic variance between different traits allowed the identification of consistent and novel patterns in the transmission of genetic variability, which may have a direct impact on selection strategies. The mean, standard deviation, and range of gametic covariance between traits were analyzed by chromosome, as illustrated in Figure 5 and Figure S1. Figure 6 presents histograms of correlation of gametic diversity between different traits for the whole population.

Traits with moderate to high correlations exhibited consistently mean gametic covariances across all chromosomal regions, always maintaining the same direction. For instance, AFC and HP showed negative gametic covariance, reflecting the strong negative correlation estimated between these traits (-0.78) (Figure 5). Similarly, the positive gametic correlation between WWd and WWm (0.72) indicates that the gametic transmission of these traits is highly associated across chromosomes (Figure 5).

In contrast, several trait combinations with low or negligible correlations exhibited a heterogeneous distribution of gametic covariance across chromosomes, with some chromosomes contributing positively and others negatively to the total covariance. This pattern was observed for pairs such as AFC x SC, HP x SC, HP x WWd, HP x WWm, among others (Figure 5). These results suggest that, for some traits, different chromosomes may exert opposing effects on the overall genetic correlation, potentially canceling their effects and resulting in lower or non-significant genome-wide correlations. Figure 5 and Figure S1 highlights that chromosomes BTA1 and BTA8 play central roles in transmitting gametic variance across growth and reproductive traits, respectively. BTA1 showed the highest variation in gametic covariance for growth traits (YW, WWd, WWm), while BTA8 exhibited the highest mean and variation in gametic covariance for reproductive traits (AFC, HP, HR).

This pattern reflects the fundamental role of genes located in these chromosomes in regulating these traits. Previous studies have demonstrated that BTA1 harbors loci associated with body growth and energy metabolism, whereas BTA8 contains genes influencing fertility and reproductive development. For instance, Morsci et al. (2006) reported QTLs in BTA1 related to weaning weight and carcass yield, supporting the importance of this chromosome for growth. Additionally, Fritz et al. (2013) and Daetwyler et al. (2014) described deleterious haplotypes in BTA8 associated with reduced fertility in dairy cattle, which could explain the higher variation observed in gametic covariance for reproductive traits.

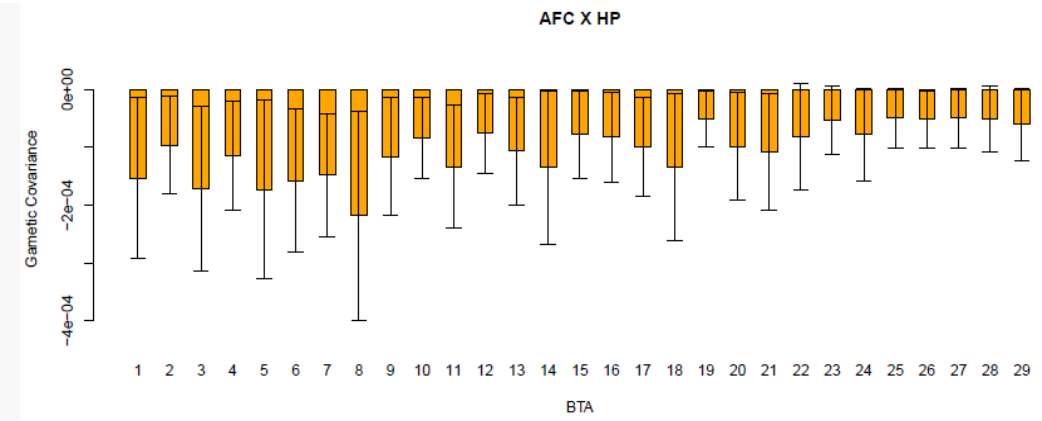
Figure 6 presents the correlations of gametic variance between traits. As expected, traits with higher traditional genetic correlations exhibited higher values of gametic variance correlation. The strong correlation between WWd and WWm (0.72) reinforces that dams transmitting greater genetic variability for direct weaning weight effects tend to do the same for maternal weaning weight effects. Conversely, the negative correlation between AFC and HP (-0.78) indicates that the gametic variability associated with these traits follows the well-established relationship between reproductive precocity and pregnancy rates.

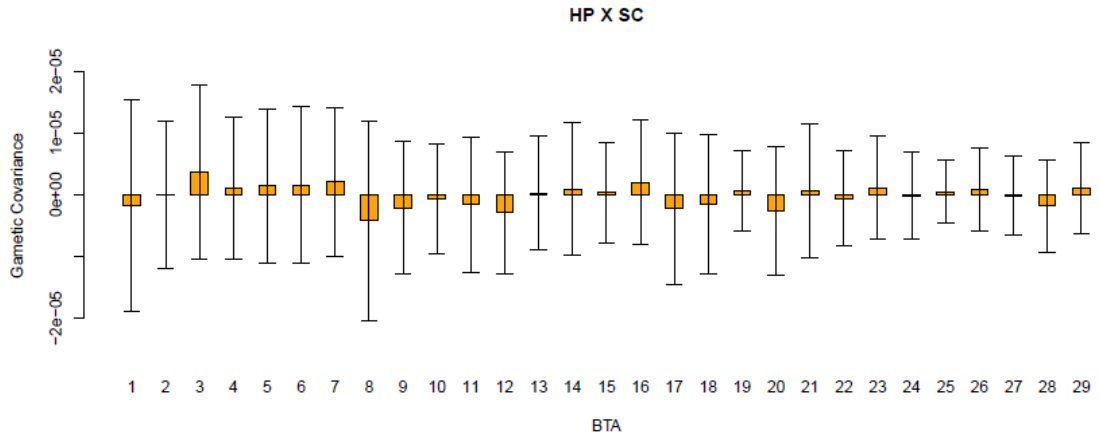
Another relevant aspect is the individual variability observed in the correlation of gametic diversity. Some trait combinations exhibited extreme values (maximum and minimum) that differed substantially from the mean,

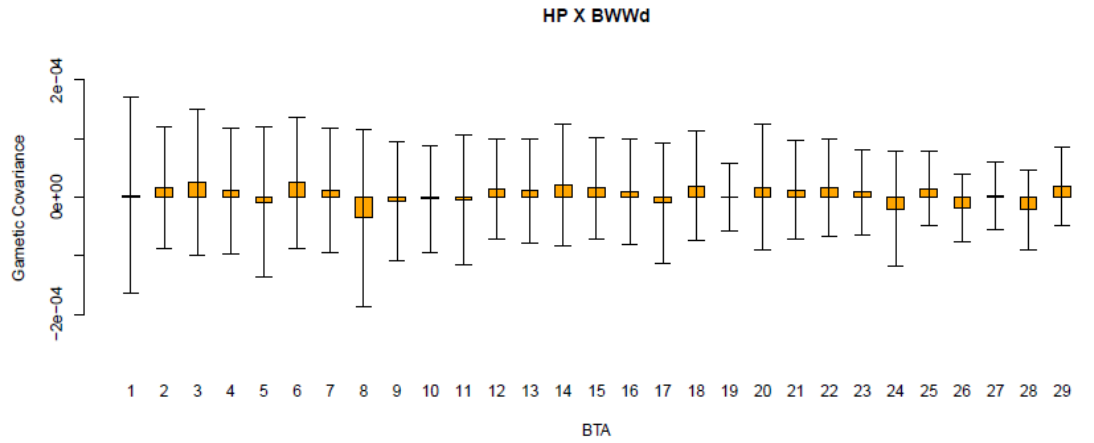
indicating that certain individuals have positive or negative gametic correlations that deviate from the population pattern. This suggests that gametic transmission can vary significantly among individuals, emphasizing the need for individualized gametic evaluation in sire selection.

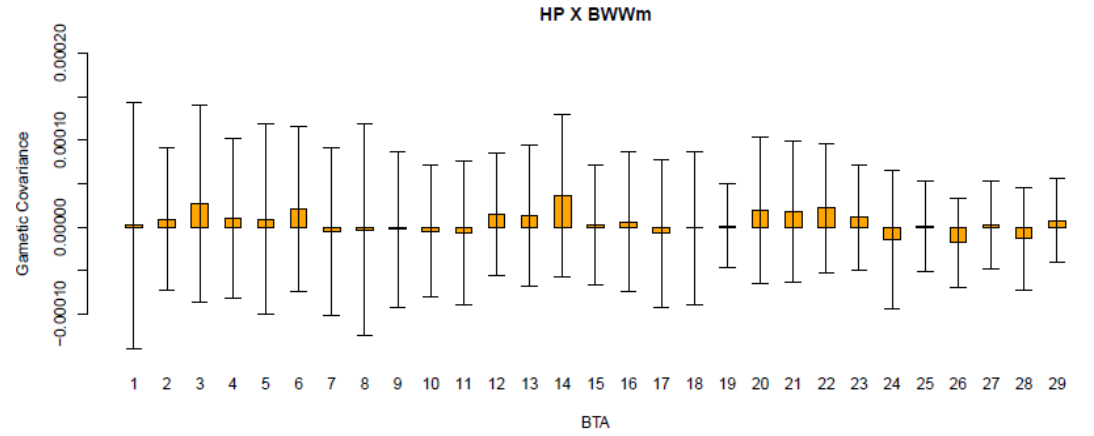
Since no prior studies have estimated the correlation of gametic variance between traits, this study provides pioneering insights by demonstrating that these correlations follow patterns like conventional genetic correlations but can be influenced by different chromosomes in a heterogeneous manner. This reinforces the importance of considering gametic variance in genetic selection, particularly for economically relevant traits.

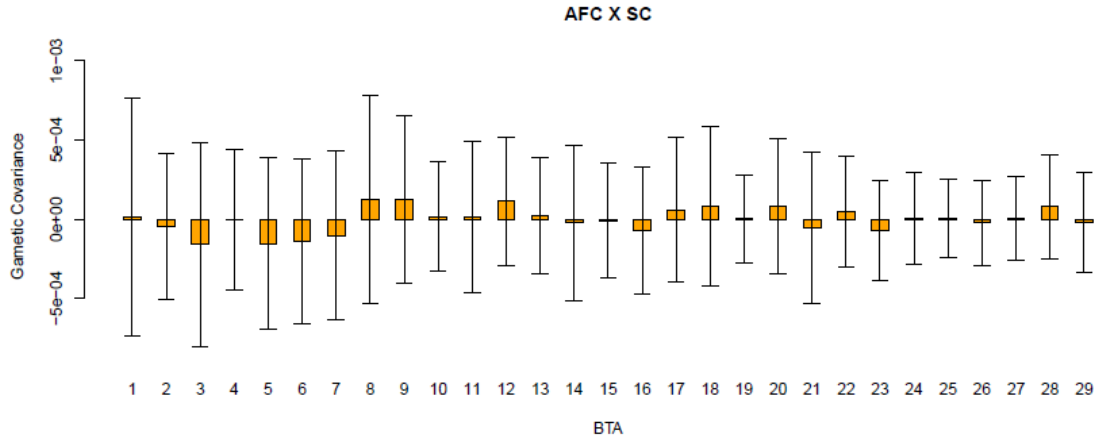
The analysis of covariance and correlation of gametic variance revealed a complex genetic architecture, in which chromosome-specific effects influence genetic correlations even at the population level. The results highlight the importance of evaluating the chromosomal structure of gametic variance, as Mendelian segregation can create distinct transmission patterns among individuals, ultimately affecting indirect selection response. Integrating gametic selection into breeding strategies has the potential to enhance multi-trait selection efficiency, accelerate genetic progress, and ensure the long-term maintenance of genetic diversity, particularly in Nellore cattle.




HP X SC
Gametic Covariance
-2e-05
0e+00
1e-05
2e-05
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA


HP X BWWd
Gametic Covariance
-2e-04
0e+00
2e-04
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA


HP X BWWm
Gametic Covariance
-0.00010
0.00000
0.00010
0.00020
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA


AFC X SC
Gametic Covariance
-5e-04
0e+00
5e-04
1e-03
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

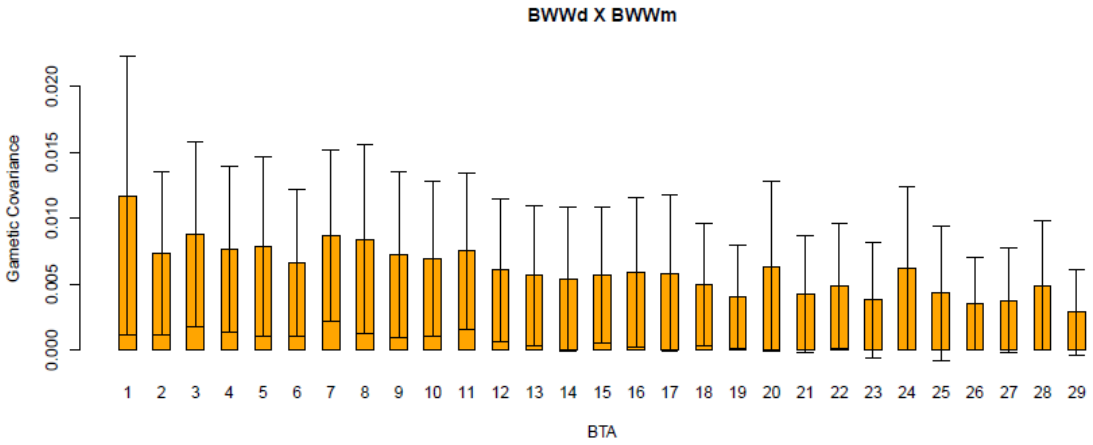


Figure 5 - Bar plots (distribution) of covariance of gametic diversity by chromosome between AFC x HP, AFC x SC, HP x SC, HP x WWd. HP x WWm and WWd x WWm. Bars indicate averages and whiskers represent standard deviations; Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)

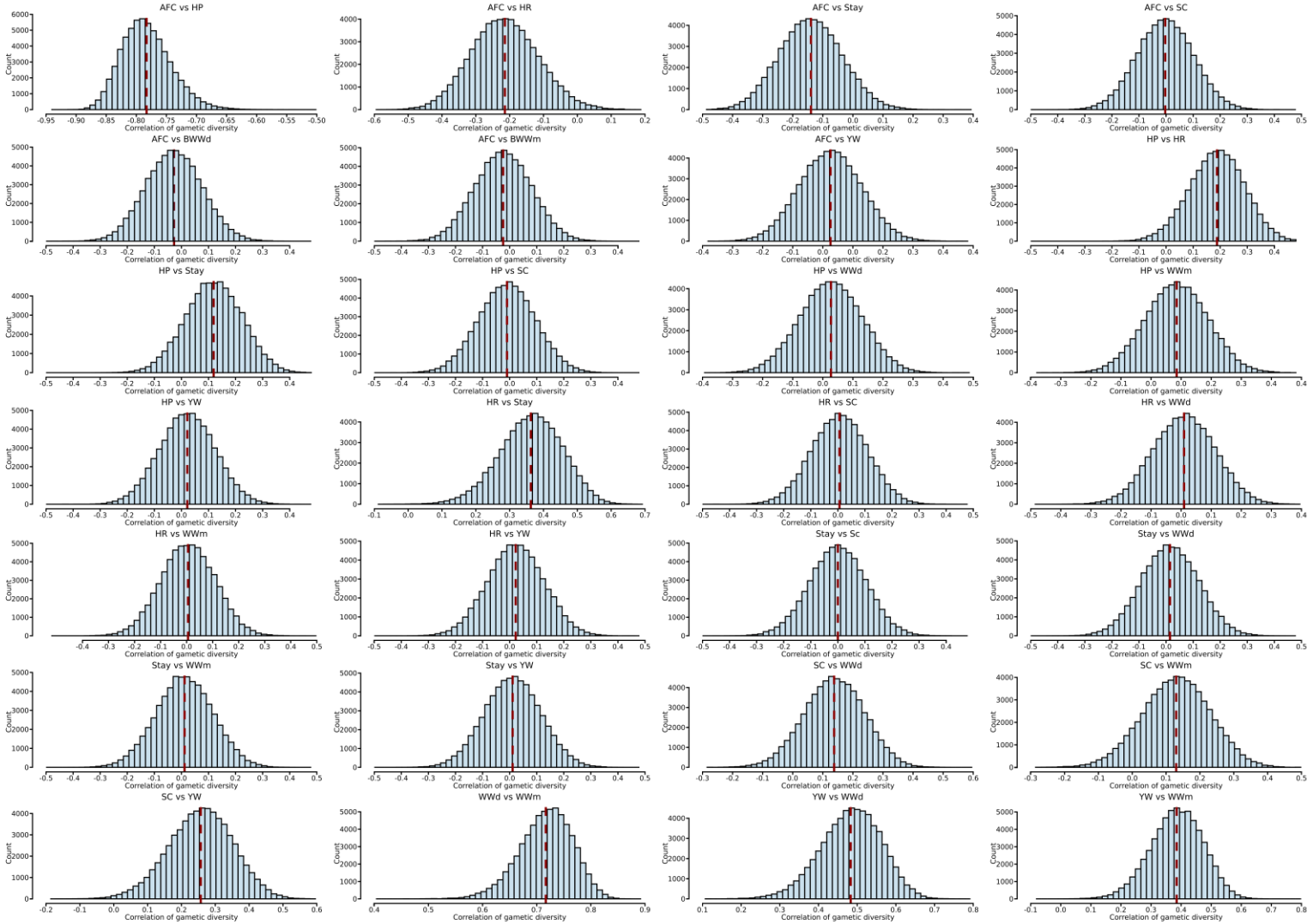


Figure 6 - Histograms of correlation of gametic diversity between different traits for the whole population; Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)

3.3.3 Application of relative indices of selection in a Nellore Population

Assessing the differential performance of relative indices of selection

The analysis of Figure 7 highlights that, assuming a gametic selection intensity of 1.5, REPD (Relative Expected Progeny Difference) exhibited greater variance, and a wider distribution compared to EPD (Estimated Progeny Difference) for most traits. This suggests that selection based on REPD may provide additional genetic gains, as this index captures not only an individual's average genetic value but also the variability in allele transmission.

However, a notable exception was observed for age at first calving (AFC), the mean EPD was higher than the REPD, which is due to the nature of this trait, as selection should always favor animals with lower values for AFC. Unlike other traits where maximizing gametic variability can be advantageous, the goal for AFC is to reduce the age at first calving, favoring individuals with a lower genetic value for this trait. For the other traits, REPD showed greater density in the tails of the distribution, indicating the presence of individuals with a higher potential to generate genetically superior offspring. This effect is particularly evident for reproductive traits such as heifer pregnancy (HP) and heifer rebreeding (HR), where incorporating gametic variance into the evaluation can enable a more efficient selection of individuals transmitting higher genetic variability for fertility. A similar pattern was observed for growth traits such as weaning weight – direct (WWd), weaning weight – maternal (WWm), and yearling weight (YW), where REPD showed greater dispersion compared to EPD. This strengthens the argument that selection based on gametic variability may enhance long-term genetic gains, especially for traits with moderate to low heritability.

As demonstrated by Santos et al. (2020), selecting a small group of sires with extreme REPD values can be an effective strategy to accelerate genetic progress. Thus, the findings in Figure 7 emphasize the importance of tailoring gametic selection to each specific trait, optimizing the use of available genetic variability without compromising production goals. This underscores REPD's potential to enhance selection criteria, balancing genetic progress and the long-term maintenance of genetic diversity in a sustainable manner.

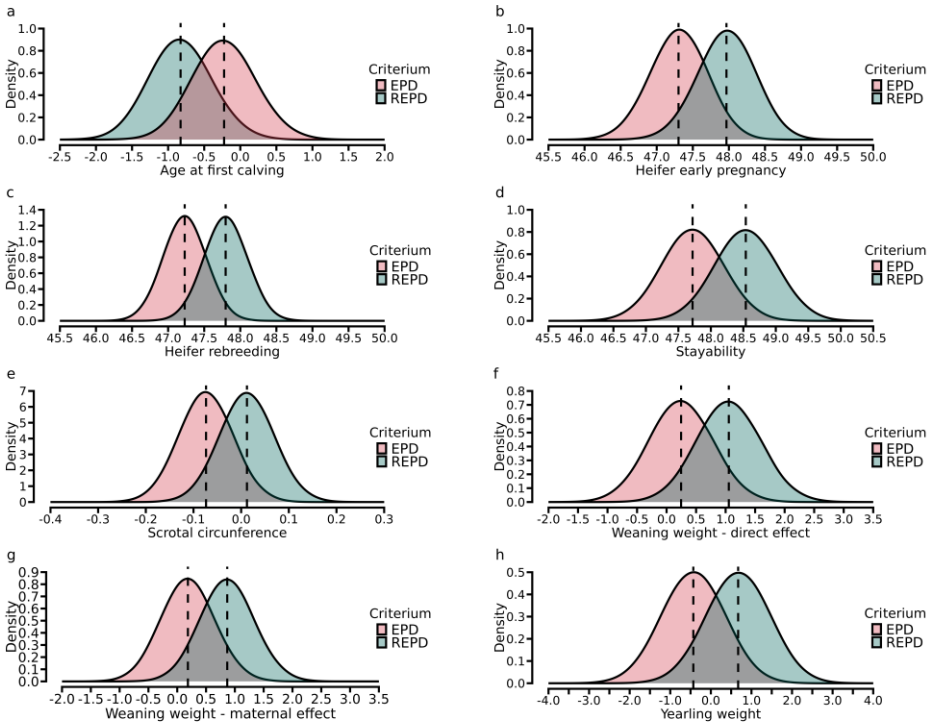


Figure 7 - Density plots of estimated genomic breeding value (EPD) and relative predicted transmitting ability (REPD) assuming future selection intensity of 1.5 for the whole population; Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)

For each trait, the ranking of the indices REPD1 and REPD1.5 was compared with EPD in different selection scenarios, including all animals, sires, and dams. In all cases, the Spearman correlation coefficients were consistently close to one (>0.99), indicating that when selecting a large proportion of animals, the ranking remains stable regardless of the index used. This confirms that the relative indices maintain the same discrimination power, direction, and orientation as traditional selection. Therefore, compared with the traditional EPD, the RPED could be considered safe index, with low risk of change the overall direction in selection, but adding some precision when a small number of animals is selected, as discussed below.

However, when analyzing the consistency of selection for different proportions of sires, the rate of coincidence decreased as the selection intensity increased (Table 5). The lowest coincidence rate (0.80) was observed when selecting only five sires, reinforcing that more stringent selection results in greater

differences between indices. Despite this decline, the overall coincidence rates remained high, demonstrating that selection using gametic variance-based indices does not drastically alter ongoing selection but rather refines it by adding precision. Among the evaluated traits, reproductive and growth-related characteristics showed the lowest coincidence rates, particularly heifer pregnancy (HP) and yearling weight (YW), with values ranging from 0.80 to 1.00 for both REPD1 and REPD1.5. This suggests that these traits may be more sensitive to changes in ranking when using gametic variance-based selection, possibly due to differences in genetic architecture and the contribution of Mendelian sampling variance.

As expected, the coincidence rates increased as the selection intensity was relaxed. When a larger number of sires was considered, the ranking remained highly consistent, confirming that the implementation of gametic variance indices does not disrupt genetic evaluation but rather provides additional resolution for identifying genetically superior individuals. These findings highlight the potential of gametic selection to optimize breeding programs by refining selection without causing abrupt changes in the current ranking system. The ability to select animals with a higher degree of genetic variability while maintaining high coincidence rates with traditional selection approaches ensures that breeding programs can integrate this methodology smoothly.

Table 5 - Percentage of selection coincidence when different subset sizes of 1978 sires ranked by traditional genetic evaluation (EPD) and by REPD indices are selected for different traits

Trait	REPD1							REPD1.5						
	5	10	50	100	200	500	1000	5	10	50	100	200	500	1000
AFC	1	0,9	0,96	0,92	0,94	0,97	0,98	1	0,9	0,92	0,9	0,92	0,95	0,96
HP	1	0,8	0,94	0,95	0,94	0,97	0,98	1	0,8	0,9	0,92	0,91	0,95	0,97
HR	0,8	0,9	0,94	0,93	0,94	0,96	0,98	0,8	0,9	0,88	0,9	0,92	0,95	0,96
STAY	1	0,9	0,9	0,9	0,93	0,96	0,98	0,8	0,9	0,88	0,85	0,89	0,95	0,96
SC	0,8	1	0,96	0,92	0,945	0,96	0,98	0,8	0,9	0,94	0,89	0,91	0,93	0,97

YW	1	0,9	0,94	0,92	0,95	0,95	0,98	0,8	0,9	0,92	0,88	0,92	0,93	0,97
WWd	0,8	0,9	0,96	0,94	0,935	0,97	0,98	0,8	0,8	0,9	0,91	0,91	0,95	0,97
WWm	0,8	0,8	0,98	0,93	0,95	0,96	0,98	0,8	0,8	0,92	0,92	0,92	0,93	0,97

Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)

Some examples of traditional plus gametic evaluation in the Nellore population are presented in a practical summary (Table 6). Here it is possible observe changes of ranking between different indices and allow us to make inference on some threshold values for each trait. Table 6 provides a comparative overview of traditional genetic evaluation and relative indices incorporating gametic variance. The analysis of the top 10 sires for each trait reveals that, despite slight changes in ranking among the different indices, the highest-ranking individuals remain well-positioned. This result indicates that incorporating gametic variance refines selection without causing drastic changes in the hierarchy of animals, which is desirable in genetic improvement programs.

An important aspect of this analysis is the probability of an individual transmitting extreme genetic values to its progeny. For age at first calving (AFC), this probability reflects the likelihood of producing offspring with lower-than-expected values, which is advantageous as selection aims to reduce this trait to improve sexual precocity. For heifer pregnancy (HP), stayability (STAY), and scrotal circumference (SC), higher values indicate a greater probability of producing progeny with superior performance, favoring the selection of genetically superior and more fertile animals. Since SC is associated with male fertility, its inclusion in the gametic analysis allows the selection of sires with greater potential to produce early and reproductively efficient offspring. For growth traits such as yearling weight (YW) and weaning weight (WWd and WWm), the estimated probability highlights the chance of a sire generating progeny with superior performance in these characteristics. This information can be valuable in selecting sires that maximize genetic gain without compromising the genetic variability of the population, allowing for more balanced and sustainable progress across generations.

Additionally, the values presented indicate the variable number of progeny (N) required to adequately capture the genetic variability of each sire. This number, estimated based on gametic variance, reflects the heterogeneity in Mendelian segregation and its influence on genetic transmission across generations. Incorporating gametic variance into genetic evaluation enhances the understanding of how genetic values are transmitted to progeny and improves the precision of identifying the best sires. These findings underscore the value of this approach in optimizing both reproductive efficiency and productive performance, ensuring that genetic selection prioritizes animals with high average genetic value as well as the potential to enhance variability within the population.

Table 6 – 10 top sires ranked by traditional genetic evaluation of different trait and their rank changes in relative indices (with gametic standard deviation weighted with 1 and 1.5). Additional parameters and inferences based on their variance of gametic diversity are also displayed

AFC										
Sire	VarG	CRV	N	EPD	Rank	REPD1	Rank	REPD1.5	Rank	Pr < -1
60705	0,13	0,43	70	-1,62	1	-1,26	1	-1,08	1	0,96
60878	0,15	0,46	80	-1,54	2	-1,15	2	-0,96	2	0,92
23823	0,16	0,47	84	-1,53	3	-1,12	3	-0,92	3	0,90
44654	0,15	0,45	77	-1,47	4	-1,08	4	-0,89	4	0,89
32	0,20	0,50	97	-1,45	5	-1,01	5	-0,79	5	0,85
28573	0,19	0,50	98	-1,40	6	-0,96	7	-0,74	9	0,82
44301	0,15	0,46	80	-1,36	7	-0,97	6	-0,77	7	0,82
61883	0,13	0,43	72	-1,33	8	-0,96	8	-0,78	6	0,81
243	0,14	0,44	75	-1,31	9	-0,94	9	-0,75	8	0,80
40	0,15	0,46	80	-1,25	10	-0,86	11	-0,66	12	0,74
HP										
Sire	VarG	CRV	N	EPD	Rank	REPD1	Rank	REPD1.5	Rank	Pr>0.02

243	7,43E-05	0,44	74	3,30E-02	1	4,16E-02	1	0,05	1	0,93
23823	8,91E-05	0,47	84	3,02E-02	2	3,97E-02	2	0,04	2	0,86
57165	8,17E-05	0,45	79	2,89E-02	3	3,79E-02	3	0,04	3	0,84
60705	6,85E-05	0,42	68	2,75E-02	4	3,57E-02	4	0,04	4	0,82
44444	8,24E-05	0,45	79	2,57E-02	5	3,47E-02	5	0,04	5	0,73
8796	9,12E-05	0,47	85	2,49E-02	6	3,44E-02	6	0,04	6	0,70
239	7,28E-05	0,43	72	2,46E-02	7	3,31E-02	11	0,04	12	0,70
3461	9,73E-05	0,48	89	2,42E-02	8	3,41E-02	7	0,04	7	0,67
43601	8,01E-05	0,45	77	2,42E-02	9	3,31E-02	10	0,04	10	0,68
44675	8,22E-05	0,45	79	2,38E-02	10	3,28E-02	12	0,04	11	0,66

HR

Sire	VarG	CRV	N	EPD	Rank	REPD1	Rank	REPD1.5	Rank	Pr>0.025
56428	5,64E-05	0,46	81	2,08E-02	1	2,84E-02	2	3,21E-02	2	0,29
57163	8,18E-05	0,53	107	2,06E-02	2	2,96E-02	1	3,42E-02	1	0,31
61842	6,20E-05	0,47	86	1,92E-02	3	2,71E-02	3	3,10E-02	3	0,23
54784	6,78E-05	0,49	92	1,84E-02	4	2,67E-02	4	3,08E-02	4	0,21
60654	6,02E-05	0,47	85	1,83E-02	5	2,61E-02	6	3,00E-02	6	0,19
43571	5,69E-05	0,46	80	1,82E-02	6	2,57E-02	7	2,95E-02	8	0,18
54444	7,07E-05	0,50	95	1,82E-02	7	2,66E-02	5	3,08E-02	5	0,21
9799	5,07E-05	0,44	74	1,64E-02	8	2,36E-02	14	2,71E-02	16	0,11
44241	8,41E-05	0,54	111	1,61E-02	9	2,53E-02	8	2,99E-02	7	0,17
61895	6,91E-05	0,49	94	1,61E-02	10	2,44E-02	10	2,86E-02	10	0,14

STAY

Sire	VarG	CRV	N	EPD	Rank	REPD1	Rank	REPD1.5	Rank	Pr>0.06
55143	1,35E-04	0,44	76	5,26E-02	1	6,42E-02	1	7,00E-02	1	0,26
44491	1,28E-04	0,43	72	4,72E-02	2	5,85E-02	2	6,41E-02	2	0,13

32	1,00E-04	0,39	59	4,63E-02	3	5,63E-02	3	6,13E-02	6	0,09
3473	1,27E-04	0,43	70	4,47E-02	4	5,60E-02	5	6,16E-02	4	0,09
55488	1,40E-04	0,45	77	4,42E-02	5	5,61E-02	4	6,20E-02	3	0,09
60980	1,43E-04	0,46	80	4,27E-02	6	5,46E-02	7	6,06E-02	9	0,07
9247	1,64E-04	0,47	87	4,22E-02	7	5,50E-02	6	6,14E-02	5	0,08
56981	1,57E-04	0,47	86	4,19E-02	8	5,44E-02	8	6,07E-02	8	0,07
9785	1,52E-04	0,47	83	4,16E-02	9	5,40E-02	10	6,01E-02	10	0,07
44400	1,49E-04	0,46	82	4,16E-02	10	5,38E-02	11	5,99E-02	12	0,07

SC

Sire	VarG	CRV	N	EPD	Rank	REPD1	Rank	REPD1.5	Rank	Pr>0.15
53134	4,19E-03	0,50	98	0,14	1	0,20	1	0,23	1	0,42
11194	3,42E-03	0,47	83	0,11	2	0,17	2	0,19	2	0,23
56811	3,11E-03	0,45	76	0,09	3	0,14	5	0,17	5	0,14
3519	2,63E-03	0,42	67	0,09	4	0,14	6	0,17	9	0,11
56713	3,77E-03	0,48	87	0,09	5	0,15	3	0,18	3	0,15
54739	3,65E-03	0,48	87	0,08	6	0,14	4	0,17	4	0,14
60675	3,49E-03	0,47	84	0,08	7	0,14	8	0,17	8	0,11
9804	4,01E-03	0,49	94	0,08	8	0,14	7	0,17	6	0,12
56295	3,96E-03	0,49	91	0,07	9	0,14	9	0,17	7	0,11
55148	3,81E-03	0,48	89	0,07	10	0,13	10	0,16	11	0,10

YW

Sire	VarG	CRV	N	EPD	Rank	REPD1	Rank	REPD1.5	Rank	Pr>2.7
53918	0,69	0,51	99	2,31	1	3,15	1	3,56	1	0,32
44626	0,56	0,46	82	1,77	2	2,51	2	2,89	2	0,11
57092	0,47	0,43	72	1,74	3	2,43	3	2,77	4	0,08
56864	0,49	0,44	74	1,70	4	2,40	5	2,75	6	0,08

43661	0,53	0,45	79	1,69	5	2,42	4	2,78	3	0,08
55114	0,59	0,48	87	1,62	6	2,39	6	2,77	5	0,08
3492	0,60	0,48	88	1,57	7	2,34	7	2,73	7	0,07
44642	0,44	0,42	69	1,53	8	2,20	12	2,53	14	0,04
44675	0,49	0,44	74	1,53	9	2,23	10	2,58	10	0,05
44689	0,53	0,45	79	1,52	10	2,25	9	2,61	9	0,05

WWd

Sire	VarG	CRV	N	EPD	Rank	REPD1	Rank	REPD1.5	Rank	Pr>2.5
44621	0,31	0,46	83	1,95	1	2,51	1	2,79	1	0,21
44638	0,34	0,48	88	1,85	2	2,43	2	2,72	2	0,17
56763	0,32	0,47	85	1,83	3	2,40	4	2,68	5	0,16
54130	0,36	0,49	93	1,80	4	2,40	3	2,70	3	0,16
43555	0,27	0,44	76	1,77	5	2,29	6	2,56	7	0,11
54655	0,27	0,44	74	1,77	6	2,29	7	2,55	9	0,11
61292	0,38	0,50	98	1,76	7	2,37	5	2,68	4	0,15
46604	0,24	0,41	65	1,75	8	2,23	10	2,47	12	0,09
56971	0,31	0,47	84	1,71	9	2,27	9	2,55	8	0,11
57206	0,26	0,43	71	1,71	10	2,22	12	2,47	13	0,09

WWm

Sire	VarG	CRV	N	EPD	Rank	REPD1	Rank	REPD1.5	Rank	Pr>2.1
54581	0,28	0,51	100	1,57	1	2,10	1	2,36	1	0,16
22906	0,22	0,47	86	1,52	2	1,99	2	2,22	2	0,11
44469	0,22	0,47	84	1,49	3	1,96	3	2,20	4	0,10
53918	0,22	0,47	85	1,49	4	1,96	5	2,19	5	0,10
61896	0,19	0,44	74	1,48	5	1,92	7	2,14	8	0,08
43700	0,21	0,45	79	1,48	6	1,93	6	2,16	6	0,08

53932	0,24	0,48	89	1,47	7	1,96	4	2,20	3	0,10
44638	0,23	0,48	87	1,42	8	1,90	8	2,14	7	0,08
61941	0,18	0,43	71	1,40	9	1,82	13	2,04	13	0,05
61814	0,14	0,39	59	1,38	10	1,76	17	1,95	22	0,03

Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)

3.3.4 Suitability of relative selection in the Nellore Population

The expected genetic gain through selection using relative indices can be estimated based on the formula proposed by Santos et al. (2019):

$$\Delta GR = r * i * \sqrt{\sigma_u^2 + 4 * var\left(\sigma_{gametic}^2\right) * i_f^2 * r_{gametic}^2}$$

where σ_u^2 represents the additive genetic variance, $var\left(\sigma_{gametic}^2\right)$ is the population variance of gametic diversity variance, i and i_f correspond to the selection intensity in the current and next generations, respectively, and r and $r_{gametic}$ represent the accuracies of genetic evaluation and gametic variance prediction. There is no a numeric population estimator for $var\left(\sigma_{gametic}^2\right)$ as we have for σ_u^2, therefore, in order to compare the traditional and relative selection gains, the variance of $\hat{\sigma}_{gametic}^2$ of the individuals was scaled through the variance of the effect of allelic substitution considering homogeneous across all loci (σ_a^2), which have independent inheritance in BLUP models, similarly to Mendelian sampling variance (*var(m)*) for one parent (Santos et al., 2019). Knowing that $var(m) = 0.25 * Nhet * \sigma_a^2$ (Santos et al., 2019), the approximation:

$$\widehat{var\left(\sigma_{gametic}^2\right)} \approx var(m)^{\frac{\log\left(var\left(\hat{\sigma}_{gametic}^2\right)\right)}{\log\left(0.25 * \sum_{h=1}^{Nhet} \alpha_h^2\right)}}$$

was used to reduce the difference in scale between the populational parameter σ_u^2 and $\widehat{var\left(\sigma_{gametic}^2\right)}$.

Considering that genetic gain by traditional selection is $\Delta G = r * i * \sigma_u$ in a scenario with $i = 1$ and $i_f = 1\ and\ 1.5$ the expected efficiency of relative selection

with respect to the traditional selection ($\Delta GR/\Delta G$) can be calculated. Disregarding the gametic accuracy, the expected gain is presented in Table 9. If $r_{gametic}$ is considered in the formula, the ΔGR is expected to be lower. Simulation studies using SNP panels suggest that gametic variance accuracy (0.50 to 0.80) is slightly lower than estimated breeding value (EBV) accuracies (Santos et al., 2019).

The results in Table 7 show that, for most traits, the estimated genetic gain under relative selection (ΔGR) was slightly higher than that obtained through traditional selection (ΔG). The increase in efficiency varied according to the trait, being more pronounced for reproductive traits (e.g., HR, STAY, SC) and maternal growth traits (WWd, WWm). The highest improvements were observed for HR (8.41%) and STAY (8.43%) under REPD1.5, suggesting that the $var(\widehat{\sigma^2_{gametic}})$ approximation could be overestimated for threshold traits, since their $var(m)$ values are very small and, therefore, sensitive to exponential adjustments.

For growth traits such as YW, WWd, and WWm, the efficiency of relative selection showed a moderate increase (2.37% to 5.38%). This can be explained by the already significant genetic progress achieved through traditional selection, reducing the potential for further improvement via gametic selection. However, for AFC, a trait where selection aims to reduce values, an increase of 2.54% under REPD1.5 reinforces the potential of gametic selection for traits under strong directional selection.

The genetic gains observed in our study, ranging from 2.37% to 8.43%, are consistent with those reported in the literature, reinforcing the effectiveness of selection based on gametic diversity variance. Santos et al. (2019) reported an increase of up to 16% over 10 generations when incorporating this approach into selection, highlighting its potential to enhance genetic progress. Bijma et al. (2020) demonstrated that adopting gametic variance could increase genetic gain by 3.6% under intense selection, a value comparable to the gains observed for growth traits in our study (2.37% to 5.38%). Additionally, Tabet (2022) showed that selection based on Mendelian sampling resulted in an accumulated gain of 8.5%, while the inclusion of gametic variance led to a 3% increase in genetic progress. Similarly, Niehoff et al. (2024a) introduced a selection criterion that

incorporated gametic variance to estimate the potential of grand offspring, resulting in a 5.8% increase in genetic gain after five generations and the preservation of 25% more genetic variability.

Although the results are positive, it is essential to consider the potential limitations of intensifying selection based on gametic diversity variance. Assigning excessively high weights to this approach may reduce the expected genetic gains, especially when selected individuals do not have enough progenies to accurately estimate their gametic variability. This is evident in Table 7, where, for some trials, the additional gain provided by REPD1.5 was relatively small compared to REPD1. This suggests that while weighting gametic variance can enhance selection, its impact may be limited when applied excessively.

Another important aspect is balancing genetic gain and maintaining genetic diversity. Selection strategies that maximize short-term genetic progress may reduce the genetic variability available for future generations, affecting the population's ability to respond to environmental changes and new breeding goals. Thus, the incorporation of gametic variance into selection criteria should be applied optimally.

Table 7 – Expected genetic gains selecting by traditional EBVs (ΔG) and relative indices (ΔGR - with gametic standard deviation weighted with 1 and 1.5). The percentage of efficiency of selection (multiplied by 100) is also presented.

Trait	$\hat{\sigma}_u^2$	$var(\widehat{\hat{\sigma}^2_{gametic}})$	ΔG	$\Delta GR1$	$\left(\frac{\Delta GR1}{\Delta G} - 1\right)\%$	$\Delta GR1.5$	$\left(\frac{\Delta GR1.5}{\Delta G} - 1\right)\%$
AFC	243,66	1,40E+00	15,610	15,788	1,140 %	16,007	2,547 %
HP	0,18	4,86E-04	0,430	0,432	0,524 %	0,435	1,175 %
HR	0,13	2,60E-03	0,365	0,379	3,820 %	0,396	8,406 %
STAY	0,27	5,29E-03	0,520	0,540	3,833 %	0,564	8,435 %
SC	2,52	4,57E-02	1,588	1,645	3,558 %	1,713	7,842 %
YW	223,59	1,19E+00	14,953	15,112	1,062 %	15,308	2,373 %
WWm	45,89	5,63E-01	6,774	6,939	2,424 %	7,138	5,376 %
WWd	45,62	5,37E-01	6,754	6,911	2,327 %	7,103	5,164 %

Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)

This study is the first to evaluate the variance of gametic diversity in the Nellore population, highlighting its influence on individual selection response and its potential to enhance breeding criteria. The results demonstrate that while traditional selection effectively identifies individuals with the highest genetic merit, incorporating gametic variance allows for differentiation of those with a greater capacity to transmit variability to their offspring, which can be particularly advantageous for traits with lower heritability. This aspect becomes even more relevant in scenarios where sires produce many descendants, such as in programs utilizing reproductive biotechnologies (AI, ET, and IVF), providing greater control over the genetic diversity being transmitted. Furthermore, considering gametic variance can contribute to more effective mating strategies in the Nellore breed, ensuring consistent genetic progress without compromising the diversity needed for future adaptation. These findings also highlight the potential of the already available genomic information and the need for its better utilization, maximizing the return on investments made in breeding programs and by producers in genotyping their herds. Integrating these metrics into genetic evaluations not only enhances selection accuracy but also strengthens the sustainability of breeding programs, making them more efficient and better prepared for future challenges.

3.4 CONCLUSION

This study was the first to estimate the variance of gametic diversity in Nellore cattle, demonstrating its influence on selection response and its relevance as a complementary criterion in genetic improvement. Chromosomal analysis revealed that BTA1 and BTA8 play a central role in transmitting variability for growth and reproductive traits, respectively. BTA1 had the greatest impact on the gametic covariance of YW, WWd, and WWm, while BTA8 was associated with AFC, HP, and HR, reinforcing the role of these chromosomes in the genetic regulation of these traits. Additionally, this was the first study to report gametic variance covariances and correlations between different traits in Nellore cattle. Moderate correlations were observed among growth traits (YW, WWd, and WWm) and among reproductive traits (HP, HR, and STAY), suggesting that the

variability transmitted to gametes is influenced by shared genomic regions. These correlations indicate that gametic diversity variance can have a combined impact on the selection of multiple traits.

The inclusion of gametic variance in selection resulted in additional genetic gains. The number of progenies required to capture 90% of the gametic variance indicated that relative selection may be particularly useful in breeding programs that employ high-prolificacy reproductive biotechnologies. Although the weighting gametic variance improved selection efficiency, excessively high weightings reduced expected gains, especially in individuals with few progenies. Thus, incorporating this criterion should be balanced to avoid losses in accuracy and genetic progress. It is evident that maximizing the use of available genomic information in Nellore breeding programs is crucial. Considering gametic variance in genetic evaluations can enhance traditional selection methods, contributing to more precise and sustainable long-term breeding strategies.

3.5 REFERENCES

ABIEC. Brazilian Livestock Profile Contents. (2023).

AGUILAR, I., MISZTAL, I., TSURUTA, S., LEGARRA, A., WANG, H. PREGSF90–POSTGSF90: computational tools for the implementation of single-step genomic selection and genome-wide association with ungenotyped individuals in BLUPF90 programs. In: **Proceedings of the 10th world congress of genetics applied to livestock production**. 2014.

BIJMA, P., WIENTJES, Y. C., & CALUS, M. P. Breeding top genotypes and accelerating response to recurrent selection by selecting parents with greater Gametic variance. **Genetics**, v. 214, n. 1, p. 91-107, 2020.

CARDOSO, D. F., DE SOUZA, F. R. P., DE CAMARGO, G. M. F., DA SILVA FONSECA, P. D., FONSECA, L. F. S., BRAZ, C. U., & TONHATI, H. Polymorphism analysis in genes of the somatotropic axis in Nellore cattle selected for growth. **Gene**, v. 545, n. 2, p. 215-219, 2014.

CARVALHEIRO, R. Genomic selection in Nelore cattle in Brazil. In: **Proceedings of the 10th World Congress on Genetics Applied to Livestock Production**. Vancouver, BC: Volume Species Breeding Beef Cattle, 2014. p. 17-22.

COLE, J. B.; VANRADEN, P. M. Use of haplotypes to estimate Mendelian sampling effects and selection limits. **Journal of Animal Breeding and Genetics**, v. 128, n. 6, p. 446-455, 2011.

COUSMINER, D. L., STERGIAKOULI, E., BERRY, D. J., ANG, W., GROEN-BLOKHUIS, M. M., KOERNER, A., & EARLY GROWTH GENETICS (EGG) CONSORTIUM. Genome-wide association study of sexual maturation in males and females highlights a role for body mass and menarche loci in male puberty. **Human molecular genetics**, v. 23, n. 16, p. 4452-4464, 2014.

DAETWYLER, H. D., CAPITAN, A., PAUSCH, H., STOTHARD, P., VAN BINSBERGEN, R., BRØNDUM, R. F. & HAYES, B. J. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. **Nature genetics**, v. 46, n. 8, p. 858-865, 2014.

DAETWYLER, H. D., HAYDEN, M. J., SPANGENBERG, G. C., & HAYES, B. J. Selection on optimal haploid value increases genetic gain and preserves more genetic diversity relative to genomic selection. **Genetics**, v. 200, n. 4, p. 1341-1348, 2015.

FERNANDES JÚNIOR, G. A., SILVA, D. A., MOTA, L. F. M., DE MELO, T. P., FONSECA, L. F. S., SILVA, D. B. D. S. & ALBUQUERQUE, L. G. Sustainable intensification of beef production in the tropics: The role of genetically improving sexual precocity of heifers. **Animals**, v. 12, n. 2, p. 174, 2022.

FRITZ, S., CAPITAN, A., DJARI, A., RODRIGUEZ, S. C., BARBAT, A., BAUR, A. & BOICHARD, D. Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2. **PloS one**, v. 8, n. 6, p. e65550, 2013.

GOIFFON, M., KUSMEC, A., WANG, L., HU, G., & SCHNABLE, P. S. Improving response in genomic selection with a population-based selection strategy: optimal population value selection. **Genetics**, v. 206, n. 3, p. 1675-1682, 2017.

KUTZLEB, C., SANDERS, G., YAMAMOTO, R., WANG, X., LICHTE, B., PETRASCH-PARWEZ, E., & KILIMANN, M. W. Paralemmin, a prenyl-palmitoyl–anchored phosphoprotein abundant in neurons and implicated in plasma membrane dynamics and cell process formation. **The Journal of cell biology**, v. 143, n. 3, p. 795-813, 1998.

LI, S., CRENSHAW III, E. B., RAWSON, E. J., SIMMONS, D. M., SWANSON, L. W., & ROSENFELD, M. G. Dwarf locus mutants lacking three pituitary cell types result from mutations in the POU-domain gene pit-1. **Nature**, v. 347, n. 6293, p. 528-533, 1990.

MEUWISSEN, T. H. E., HAYES, B. J., & GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, n. 4, p. 1819-1829, 2001.

MICHENET, A., SAINTILAN, R., VENOT, E., & PHOCAS, F. Insights into the genetic variation of maternal behavior and suckling performance of continental beef cows. **Genetics Selection Evolution**, v. 48, p. 1-12, 2016.

MILLS, M. C., TROPF, F. C., BRAZEL, D. M., VAN ZUYDAM, N., VAEZ, A., PERS, T. H. & DAY, F. R. Identification of 370 genetic loci for age at first sex and birth linked to externalising behaviour. **bioRxiv**, 2021.

MISZTAL, I., GIANOLA, D., FOULLEY, J. L. Computing aspects of a nonlinear method of sire evaluation for categorical data. **Journal of Dairy Science**, v. 72, n. 6, p. 1557-1568, 1989.

MISZTAL, I. et al. BLUPF90 and related programs (BGF90). In: **Proceedings of the 7th world congress on genetics applied to livestock production**. 2002. p. 743-744.

MORSCI, N. S.; SCHNABEL, R. D.; TAYLOR, J. F. Association analysis of adiponectin and somatostatin polymorphisms on BTA1 with growth and carcass traits in Angus cattle. **Animal Genetics**, v. 37, n. 6, p. 554-562, 2006.

MÜLLER, D., SCHOPP, P., & MELCHINGER, A. E. Selection on expected maximum haploid breeding values can increase genetic gain in recurrent genomic selection. **G3: Genes, Genomes, Genetics**, v. 8, n. 4, p. 1173-1181, 2018.

NIEHOFF, T. A., TEN NAPEL, J., BIJMA, P., POOK, T., WIENTJES, Y. C., HEGEDŰS, B., & CALUS, M. P. Improving selection decisions with mating information by accounting for Mendelian sampling variances looking two generations ahead. **Genetics Selection Evolution**, v. 56, n. 1, p. 41, 2024a.

NIEHOFF, T. A., TEN NAPEL, J., & CALUS, M. P. Prediction of additive genetic variances of descendants for complex families based on Mendelian sampling variances. **G3: Genes, Genomes, Genetics**, p. jkae205, 2024b.

NEVES, H. H., VARGAS, G., BRITO, L. F., SCHENKEL, F. S., ALBUQUERQUE, L. G., & CARVALHEIRO, R. Genetic and genomic analyses of testicular hypoplasia in Nellore cattle. **PLoS One**, v. 14, n. 1, p. e0211159, 2019.

PERRY, J. R., STOLK, L., FRANCESCHINI, N., LUNETTA, K. L., ZHAI, G., MCARDLE, P. F. & MURABITO, J. M. Meta-analysis of genome-wide association data identifies two loci influencing age at menarche. **Nature genetics**, v. 41, n. 6, p. 648-650, 2009.

PFÄFFLE, R. W., DIMATTIA, G. E., PARKS, J. S., BROWN, M. R., WIT, J. M., JANSEN, M. & INGRAHAM, H. A. Mutation of the POU-specific domain of Pit-1 and hypopituitarism without pituitary hypoplasia. **Science**, v. 257, n. 5073, p. 1118-1121, 1992.

SANTOS, D. J. A., COLE, J. B., LAWLOR JR, T. J., VANRADEN, P. M., TONHATI, H., & MA, L. Variance of gametic diversity and its application in selection programs. **Journal of dairy science**, v. 102, n. 6, p. 5279-5294, 2019.

SANTOS, D. J. A., COLE, J. B., LIU, G. E., VANRADEN, P. M., & MA, L. Gamevar. f90: a software package for calculating individual gametic diversity. **BMC bioinformatics**, v. 21, n. 1, p. 1-5, 2020.

SARGOLZAEI, M., CHESNAIS, J. P., & SCHENKEL, F. S. A new approach for efficient genotype imputation using information from relatives. **BMC genomics**, v. 15, n. 1, p. 478, 2014.

SCHAEFFER, L. R. Strategy for applying genome-wide selection in dairy cattle. **Journal of animal Breeding and genetics**, v. 123, n. 4, p. 218-223, 2006.

SEGELKE, D., REINHARDT, F., LIU, Z., & THALLER, G. Prediction of expected genetic variation within groups of offspring for innovative mating schemes. **Genetics Selection Evolution**, v. 46, n. 1, p. 42, 2014.

SHEN B., JIANG J., SEROUSSI E., LIU G. E. & MA L. Characterization of recombination features and the genetic basis in multiple cattle breeds. **BMC Genomics**, 19:304, 2018.

SORENSEN, D., & GIANOLA, D. Likelihood, Bayesian and MCMC methods in quantitative genetics. 2002.

SUZU, S., HAYASHI, Y., HARUMI, T., NOMAGUCHI, K., YAMADA, M., HAYASAWA, H., & MOTOYOSHI, K. Molecular cloning of a novel immunoglobulin superfamily gene preferentially expressed by brain and testis. **Biochemical and biophysical research communications**, v. 296, n. 5, p. 1215-1221, 2002.

TAHIR, M. S., PORTO-NETO, L. R., GONDRO, C., SHITTU, O. B., WOCKNER, K., TAN, A. W., ... & FORTES, M. R. Meta-analysis of heifer traits identified reproductive pathways in Bos indicus cattle. **Genes**, v. 12, n. 5, p. 768, 2021.

TEIXEIRA, D., ALVES FERNANDES JÚNIOR, G., BERALDO DOS SANTOS SILVA, D., BERMAL COSTA, R., TAKADA, L., GUSTAVO MANSAN GORDO, D., ... & GALVÃO DE ALBUQUERQUE, L. Genomic analysis of stayability in Nellore cattle. **PLoS One**, v. 12, n. 6, p. e0179076, 2017.

TURNER, M. D.; WILDE, C. J.; BURGOYNE, R. D. Exocytosis from permeabilized lactating mouse mammary epithelial cells. Stimulation by Ca2+ and phorbol ester, but inhibition of regulated exocytosis by guanosine 5′-[γ-thio] triphosphate. **Biochemical Journal**, v. 286, n. 1, p. 13-15, 1992.

VANRADEN, P. 2015. conflict.f90. http://aipl.arsusda.gov/software/genosim

WANG, J., FAN, H. C., BEHR, B., & QUAKE, S. R. Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. **Cell**, v. 150, n. 2, p. 402-412, 2012.

4 FINAL CONSIDERATIONS

Future research should prioritize maximizing the use of available genomic information, leveraging its full potential to refine the estimation of gametic variance and improve selection strategies. This study provided unprecedented insights into the genetic mechanisms underlying recombination and the variability of gametic diversity in Nellore cattle, a key breed for Brazilian livestock production. The findings contribute to a deeper understanding of meiotic recombination patterns, identifying key genomic regions involved in genetic variability transmission. Moreover, this research presents a novel approach to incorporating gametic variance into selection strategies, demonstrating its potential to enhance genetic gain while maintaining genetic diversity. By making available the first recombination map for the Nellore breed using a SNP panel for further studies, this work lays the foundation for future advancements in genomic selection, contributing to more sustainable and efficient breeding programs.

5 SUPPLEMENTARY MATERIAL - Characterization of recombination events and genome-wide association in Nellore males

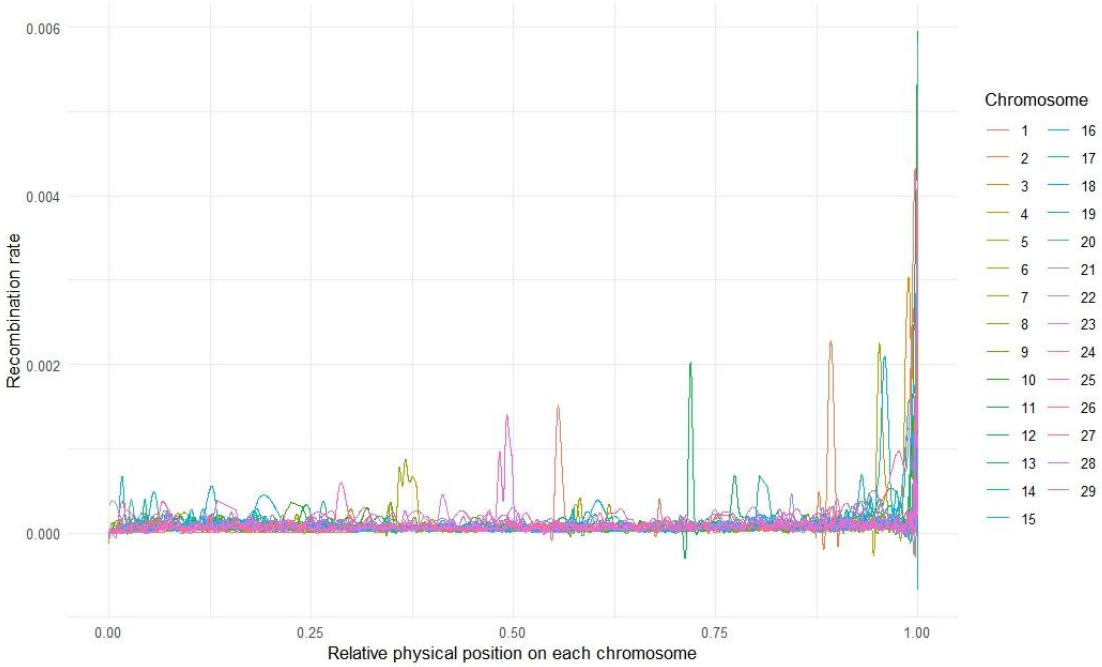


Figure S1 - The relative physical position for each SNP interval on a chromosome was calculated by standardizing the original physical position by the chromosome length: a value of zero corresponds to the beginning of a chromosome and a value of one corresponds to the end. The smooth spline model fitted across all the 29 autosomes

Table S1 – List of hotspots as SNP intervals with recombination rate >2.5 standard deviations greater than the mean (Annex 1)

Table S2 – Recombination map for Nellore animals (Annex 2)

Table S3 – Chromosome (CHR), SNP position in bp and proportion of direct additive genetic variance explained by significant SNPs (% varG) for recombination number of Nellore males

Model a				
CHR	**SNP position**	**SNP name**	**-Log$_{10}$(p-value)**	**% varG**
6	92144231	BovineHD0600026029	6.431	0.172
27	31039126	BovineHD2700008525	6.240	0.212
1	129596073	BovineHD0100037135	6.179	0.478
12	73317927	BovineHD1200021803	5.662	0.203

6	83291498	BovineHD0600023379	5.485	0.013
6	83310622	BovineHD0600023388	5.485	0.013
6	83322740	BovineHD0600023391	5.485	0.013
6	83302827	BovineHD0600023386	5.485	0.013
6	83360451	BovineHD0600023402	5.485	0.013
6	92794540	BovineHD0600026201	5.420	0.193
2	101630783	BovineHD0200029333	5.419	0.260
5	78895966	ARS-BFGL-NGS-21874	5.318	0.107
5	78901882	BovineHD0500022497	5.318	0.107
5	78904476	BovineHD0500022499	5.318	0.107
5	78903899	BovineHD0500022498	5.318	0.107
5	78907148	BovineHD0500022500	5.318	0.107
5	14907432	BovineHD0500004405	5.300	0.198
7	22053836	BovineHD0700006404	5.296	0.164
7	24931295	BovineHD0700007233	5.276	0.063
7	105544477	BovineHD0700031503	5.235	0.083
6	92813658	BovineHD0600026204	5.207	0.147
12	82795910	BovineHD1200025243	5.202	0.152
15	67362128	BovineHD1500019681	5.131	0.056
12	67312783	BovineHD1200018508	5.118	0.101
16	2350416	BovineHD4100012271	5.090	0.157
14	75886344	BovineHD1400021910	5.076	0.017
12	79277555	BovineHD1200024037	5.076	0.086
7	6395213	BovineHD4100005726	5.067	0.103
6	68099345	BovineHD0600019279	5.034	0.099

Model b

CHR	SNP position	SNP name	-Log$_{10}$(p-value)	% varG
29	34841469	BovineHD2900010694	8.945	0.210
29	34842171	BovineHD2900010695	8.945	0.210
29	34842922	BovineHD2900010696	7.258	0.033
1	129596073	BovineHD0100037135	7.092	0.510
10	47901861	BovineHD1000014419	6.907	0.257
1	128938526	BovineHD0100036946	6.796	0.027
1	128507155	BovineHD0100036703	6.792	0.058
1	128800455	BovineHD0100036867	6.759	0.087
1	128768995	BovineHD0100036845	6.746	0.099
1	128080940	BovineHD0100036510	6.734	0.110
1	128079633	BovineHD0100036509	6.734	0.110
1	128626738	BovineHD0100036753	6.702	0.009
1	129569312	BovineHD0100037115	6.576	0.040
10	47833468	BovineHD1000014404	6.529	0.133
10	47445621	BovineHD1000014276	6.496	0.021
10	47934450	BovineHD1000014430	6.480	0.064
1	128968345	BovineHD0100036959	6.455	0.015
10	47888491	BovineHD1000014415	6.440	0.173
10	47890865	BovineHD1000014417	6.440	0.173
10	47891805	ARS-BFGL-NGS-64260	6.440	0.173
6	92144231	BovineHD0600026029	6.431	0.152

1	128645157	BovineHD0100036763	6.419	0.029
1	128645858	BovineHD0100036764	6.419	0.029
1	128878546	BovineHD0100036913	6.357	0.054
1	129546015	BovineHD0100037095	6.349	0.029
1	129550873	BovineHD0100037101	6.349	0.029
10	47151345	BovineHD1000014160	6.336	0.157
1	128505255	BovineHD0100036701	6.335	0.037
29	34936540	BovineHD2900010731	6.295	0.022
10	45794472	BovineHD1000030643	6.277	0.145
10	45898922	BovineHD1000013784	6.268	0.086
10	45905824	BovineHD1000013786	6.268	0.086
10	45917202	BovineHD1000013789	6.268	0.085
27	31039126	BovineHD2700008525	6.240	0.144
1	128037217	BovineHD0100036499	6.237	0.103
10	47457525	BovineHD1000014278	6.224	0.020
1	128697052	BovineHD0100036793	6.222	0.070
1	128352710	BovineHD0100036629	6.217	0.066
1	128608694	BovineHD0100036744	6.208	0.101
1	129162836	BovineHD0100037004	6.182	0.090
1	129574633	BovineHD0100037116	6.175	0.098
1	128755742	BovineHD0100036838	6.174	0.056
1	128754794	BovineHD0100036837	6.174	0.056
1	128982676	BovineHD0100036962	6.160	0.007
1	128842936	BovineHD0100036891	6.146	0.062
1	128785022	BovineHD0100036856	6.146	0.079
1	128510430	BovineHD0100036705	6.134	0.007
1	128408335	BovineHD0100036654	6.106	0.028
29	34840886	BovineHD4100019023	6.105	0.001
1	128323943	BovineHD4100000732	6.098	0.034
1	128324703	BovineHD0100036617	6.098	0.034
1	128325557	BovineHD0100036618	6.098	0.034
10	45897958	BovineHD1000013783	6.096	0.076
1	129579717	BovineHD0100037118	6.041	0.100
1	128929146	BovineHD0100036941	6.029	0.001
29	34938498	BovineHD2900010732	6.029	0.016
1	128656901	BovineHD0100036769	5.971	0.019
10	46502211	BovineHD1000013932	5.960	0.110
1	128315979	BovineHD0100036612	5.954	0.043
29	34932613	BovineHD2900010728	5.939	0.003
1	128374395	BovineHD0100036640	5.909	0.033
1	128783960	BovineHD4100000737	5.907	0.012
1	128788137	BovineHD0100036858	5.907	0.012
1	128344162	BovineHD0100036625	5.893	0.054
1	128875411	BovineHD0100036910	5.857	0.040
1	128976215	BovineHD0100036961	5.835	0.002
10	46485339	BovineHD1000013929	5.826	0.103
1	128139303	BovineHD0100036535	5.779	0.020
1	129187047	BovineHD0100037007	5.778	0.052

1	128737460	BovineHD0100036819	5.757	0.006
1	128738150	BovineHD0100036820	5.757	0.006
1	128739175	BovineHD0100036821	5.757	0.006
1	128921870	BovineHD0100036937	5.738	0.012
1	128612913	BovineHD0100036747	5.731	0.112
1	128971583	BovineHD0100036960	5.725	0.021
1	128285508	BovineHD0100036597	5.690	0.029
29	34887929	BovineHD2900010712	5.683	0.000
1	128949907	BovineHD0100036955	5.677	0.050
10	45958934	BovineHD1000013802	5.671	0.100
10	45957875	BovineHD1000013801	5.671	0.100
12	73317927	BovineHD1200021803	5.662	0.109
1	129590492	BovineHD0100037128	5.648	0.163
1	128616588	BovineHD0100036748	5.635	0.067
1	128888895	BovineHD0100036920	5.592	0.041
1	128429011	BovineHD0100036663	5.566	0.001
29	34933724	BovineHD2900010729	5.552	0.001
1	128134839	BovineHD0100036533	5.547	0.030
1	128725264	BovineHD0100036812	5.520	0.001
29	34894802	BovineHD2900010714	5.503	0.002
29	34880258	ARS-BFGL-NGS-23717	5.489	0.005
6	83291498	BovineHD0600023379	5.485	0.026
6	83310622	BovineHD0600023388	5.485	0.026
6	83322740	BovineHD0600023391	5.485	0.026
6	83302827	BovineHD0600023386	5.485	0.026
6	83360451	BovineHD0600023402	5.485	0.026
1	128687619	BovineHD0100036785	5.463	0.012
10	45217561	BovineHD1000013593	5.458	0.101
10	45633163	BTB-01891199	5.456	0.041
10	45644611	BovineHD1000013725	5.456	0.041
10	45654705	BovineHD1000013727	5.456	0.042
10	45669628	BovineHD1000013731	5.456	0.042
10	45680616	BTB-00420673	5.456	0.042
10	45683948	BovineHD1000013737	5.456	0.042
10	45693931	BovineHD1000013740	5.456	0.042
10	45698635	BovineHD1000013741	5.456	0.042
10	45700666	BovineHD1000013743	5.456	0.042
10	45703547	BovineHD1000013744	5.456	0.042
10	45709283	BTB-00420603	5.456	0.042
10	45710996	BovineHD1000013747	5.456	0.042
10	45746903	BovineHD1000013755	5.456	0.042
10	45764666	BovineHD1000013758	5.456	0.042
10	45765993	BovineHD1000013759	5.456	0.042
10	45800362	BovineHD1000013764	5.456	0.042
10	45806574	BovineHD1000013766	5.456	0.042
10	45807766	ARS-BFGL-NGS-43266	5.456	0.042
10	45904593	BovineHD1000013785	5.456	0.042
10	45947692	BovineHD1000013793	5.456	0.042

10	45639992	BovineHD1000013724	5.456	0.041
10	45677319	BovineHD1000013733	5.456	0.042
10	45677974	BovineHD1000013734	5.456	0.042
10	45679239	BovineHD1000013735	5.456	0.042
10	45687468	BovineHD1000013738	5.456	0.042
10	45699145	BovineHD1000013742	5.456	0.042
10	45706103	BovineHD1000013745	5.456	0.042
10	45734234	BovineHD1000013753	5.456	0.042
10	45743173	BovineHD1000013754	5.456	0.042
10	45755184	BovineHD1000013757	5.456	0.042
10	45769339	BovineHD1000013760	5.456	0.042
10	45777873	BTA-67227-no-rs	5.456	0.042
10	45803672	BovineHD1000013765	5.456	0.041
1	128734264	BovineHD0100036817	5.451	0.019
1	128770665	BovineHD0100036847	5.430	0.093
1	129558037	BovineHD0100037107	5.425	0.054
1	128940245	BovineHD0100036948	5.425	0.013
6	92794540	BovineHD0600026201	5.420	0.169
1	129305891	BovineHD0100037027	5.419	0.006
2	101630783	BovineHD0200029333	5.419	0.161
1	128129098	BovineHD0100036532	5.405	0.022
29	34905460	BovineHD2900015475	5.405	0.007
29	34931380	BovineHD2900010727	5.375	0.018
1	128785863	BovineHD0100036857	5.349	0.003
29	34861752	BovineHD2900010704	5.335	0.008
1	128385204	BovineHD0100036644	5.332	0.023
1	129269731	BovineHD0100037022	5.324	0.065
1	128790942	BovineHD0100036859	5.321	0.061
1	128795972	BovineHD0100036863	5.321	0.061
5	78895966	ARS-BFGL-NGS-21874	5.318	0.118
5	78901882	BovineHD0500022497	5.318	0.118
5	78904476	BovineHD0500022499	5.318	0.118
5	78903899	BovineHD0500022498	5.318	0.118
5	78907148	BovineHD0500022500	5.318	0.118
5	14907432	BovineHD0500004405	5.300	0.173
7	22053836	BovineHD0700006404	5.296	0.160
10	47871076	Hapmap36641-SCAFFOLD21136_2257	5.295	0.037
10	45250412	BovineHD1000013599	5.292	0.081
7	24931295	BovineHD0700007233	5.276	0.038
7	105544477	BovineHD0700031503	5.235	0.074
1	128437663	BovineHD0100036669	5.231	0.007
6	92813658	BovineHD0600026204	5.207	0.118
12	82795910	BovineHD1200025243	5.202	0.141
1	129018668	BovineHD0100036971	5.191	0.015
1	128490803	BovineHD0100036694	5.168	0.003
10	45197910	BovineHD1000013590	5.166	0.078
1	128138401	BovineHD0100036534	5.138	0.063
15	67362128	BovineHD1500019681	5.131	0.065

1	128745600	BovineHD0100036826	5.130	0.022
1	128747823	BovineHD0100036829	5.130	0.022
1	128104424	BovineHD0100036523	5.126	0.027
12	67312783	BovineHD1200018508	5.118	0.076
29	34860156	BovineHD4100019024	5.116	0.003
29	34844114	BovineHD2900010697	5.092	0.014
1	129594768	BovineHD0100037133	5.091	0.020
1	129586737	BovineHD0100037124	5.091	0.020
16	2350416	BovineHD4100012271	5.090	0.075
10	45852782	BovineHD1000013774	5.083	0.030
10	45873770	BovineHD1000013779	5.083	0.030
10	45884962	ARS-BFGL-NGS-61625	5.083	0.030
10	45870583	BovineHD1000013778	5.083	0.030
10	45880714	BovineHD1000013780	5.083	0.030
10	45888410	BovineHD1000013781	5.083	0.031
1	129556062	BovineHD0100037105	5.077	0.012
12	79277555	BovineHD1200024037	5.076	0.057
14	75886344	BovineHD1400021910	5.076	0.019
1	128184812	ARS-BFGL-NGS-110636	5.071	0.031
7	6395213	BovineHD4100005726	5.067	0.081
10	45832388	BovineHD1000013770	5.058	0.034
1	128907137	BovineHD0100036931	5.050	0.004
1	128811490	BovineHD0100036875	5.043	0.010
6	68099345	BovineHD0600019279	5.034	0.095
29	34835279	BovineHD2900010693	5.021	0.007
29	34827355	BovineHD2900010691	5.011	0.009
1	128678487	ARS-BFGL-NGS-80200	5.001	0.005

Table S4 – Chromosome (CHR), SNP position in bp and proportion of direct additive genetic variance explained by significant SNPs (% varG) for recombination hotspots number of Nellore males

Model a				
CHR	**SNP position**	**SNP name**	**-Log$_{10}$(p-value)**	**% varG**
10	21086459	BovineHD1000006791	10,86366	0,014
1	157698749	BovineHD2400007403	10,57469	0,054
1	157699483	BovineHD2400007402	10,57469	0,054
1	157160634	BovineHD0100012743	10,15385	0,016
10	21079914	BovineHD1000006789	9,392016	0,012
13	58210244	BovineHD1300016819	8,750904	0,033
1	156948236	BovineHD0100039627	8,623643	0,024
13	58245472	BovineHD1300016827	8,469698	0,023
1	157673016	BovineHD0100012909	8,46946	0,051
1	157680911	BovineHD0100012913	8,46946	0,051
1	157683452	BovineHD0100012915	8,469459	0,051
1	157199865	BovineHD0100012754	8,340786	0,014
1	157188322	BovineHD0100012748	8,340785	0,014
13	58232306	BovineHD1300016825	8,07351	0,032
1	157527295	BovineHD3300000086	8,052369	0,071
1	157530630	BovineHD3300000081	8,052369	0,071
1	157532893	BovineHD3300000075	8,052369	0,071
1	157561069	BovineHD0100012862	8,052369	0,071
1	157530606	BovineHD3300000082	8,052361	0,071
1	157532486	BovineHD3300000076	8,052361	0,071
1	157532366	BovineHD3300000078	7,866561	0,032
10	21030895	BovineHD1000006768	7,132025	0,023
1	156942192	BovineHD0100039622	7,10888	0,032
1	156946290	BovineHD0100039626	7,058582	0,008
10	20993107	BovineHD1000006757	6,989711	0,029
1	156949184	BovineHD0100039628	6,506618	0,005
13	58316230	BovineHD1300016842	6,324664	0,001
10	20987470	BovineHD1000006754	6,271982	0,008
10	20958739	ARS-BFGL-BAC-10972	5,802796	0,021
10	20959530	BovineHD1000006748	5,802796	0,021
26	31375550	BovineHD2600008479	5,787662	0,016
3	90476782	BovineHD0300026193	5,514447	0,027
3	90481941	BovineHD0300026196	5,514446	0,027
13	59006701	BovineHD1300017060	5,368911	0,017
1	156902670	BovineHD0100039615	5,344414	0,033
1	157295166	BovineHD0100012800	5,241085	0,005
3	87460288	BovineHD0300025210	5,240839	0,032
1	94348223	BovineHD0100026972	5,226435	0,006
1	157348202	BovineHD0100012817	5,18989	0,004
15	33282796	BovineHD1500009157	5,176747	0,012
11	40264225	BovineHD1100011805	5,129552	0,014
11	40272630	Hapmap50157-BTA-91990	5,129551	0,014
12	58494271	BovineHD1200016164	5,104063	0,007

1	156943013	BovineHD0100039623	5,030634	0,026
		Model b		
CHR	**SNP position**	**SNP name**	**-Log$_{10}$(p-value)**	**% varG**
1	157698749	BovineHD2400007403	9,907177	0,042219
1	157699483	BovineHD2400007402	9,907177	0,042219
1	157527295	BovineHD3300000086	9,556203	0,015829
1	157530630	BovineHD3300000081	9,556203	0,015829
1	157532893	BovineHD3300000075	9,556203	0,015829
1	157561069	BovineHD0100012862	9,556203	0,015829
1	157530606	BovineHD3300000082	9,556199	0,015829
1	157532486	BovineHD3300000076	9,556199	0,015829
26	31375550	BovineHD2600008479	9,023533	0,011506
1	157673016	BovineHD0100012909	8,616489	0,038574
1	157680911	BovineHD0100012913	8,616489	0,038574
1	157683452	BovineHD0100012915	8,616487	0,038574
26	31339399	BovineHD2600008461	7,754012	0,01509
10	21957034	Hapmap41917-BTA-61237	6,856112	0,037371
10	21954752	BovineHD1000007057	6,856111	0,037371
1	157160634	BovineHD0100012743	6,819303	0,015758
26	31054322	BovineHD2600008361	6,771027	0,021885
26	31362767	BovineHD2600008471	6,560912	0,016846
10	21416197	BovineHD1000006925	6,488786	0,024011
3	85936852	BovineHD0300024704	6,145073	0,039415
9	70754671	BovineHD0900019903	6,042195	0,03828
9	70749414	BovineHD0900019899	6,042193	0,03828
9	70750941	BovineHD0900019900	6,042193	0,03828
10	21422639	BovineHD1000006927	5,739677	0,023204
9	70670645	BovineHD0900019882	5,715507	0,037666
15	33282796	BovineHD1500009157	5,676334	0,015555
26	31013105	BovineHD2600008349	5,530935	0,014322
26	31016497	BovineHD2600008350	5,530935	0,014322
26	31451577	BovineHD2600008500	5,478664	0,009353
26	31463831	BovineHD2600008502	5,478664	0,009353
28	10010697	BovineHD2800003074	5,401753	0,044886
28	9999877	BovineHD2800003072	5,401752	0,044886
26	31564827	BovineHD2600008511	5,358149	0,009213
26	31578147	BovineHD2600008512	5,358142	0,009213
3	87460288	BovineHD0300025210	5,335034	0,012525
10	21030895	BovineHD1000006768	5,325952	0,029088
6	89408744	BovineHD0600024973	5,270929	0,007404
10	21086459	BovineHD1000006791	5,254785	0,028678
21	12043366	BovineHD2100003078	5,248379	0,031185
1	157199865	BovineHD0100012754	5,244245	0,014162
1	157188322	BovineHD0100012748	5,244242	0,014162
3	84673191	BovineHD0300024344	5,217932	0,027212
8	36863443	BovineHD0800011048	5,212022	0,010321
21	12046044	BovineHD2100003081	5,121348	0,027744
10	21079914	BovineHD1000006789	5,119991	0,027912

4	104825199	BovineHD0400029833	5,082146	0,042071
14	47159178	BovineHD1400014041	5,073638	0,039144
14	47160685	BovineHD1400014042	5,073638	0,039144
14	47163738	BovineHD1400014043	5,073638	0,039144
17	63325268	BovineHD1700018883	5,071574	0,038095
10	21535687	BovineHD1000031169	5,035327	0,028812
1	157532366	BovineHD3300000078	5,010777	0,015076

6 SUPPLEMENTARY MATERIAL - Variance of gametic diversity estimation for economically important traits of Nellore cattle

Table S1 - Descriptive statistics of variance of gametic diversity and CRV (Coefficient of Relative Variation) in the Nellore population

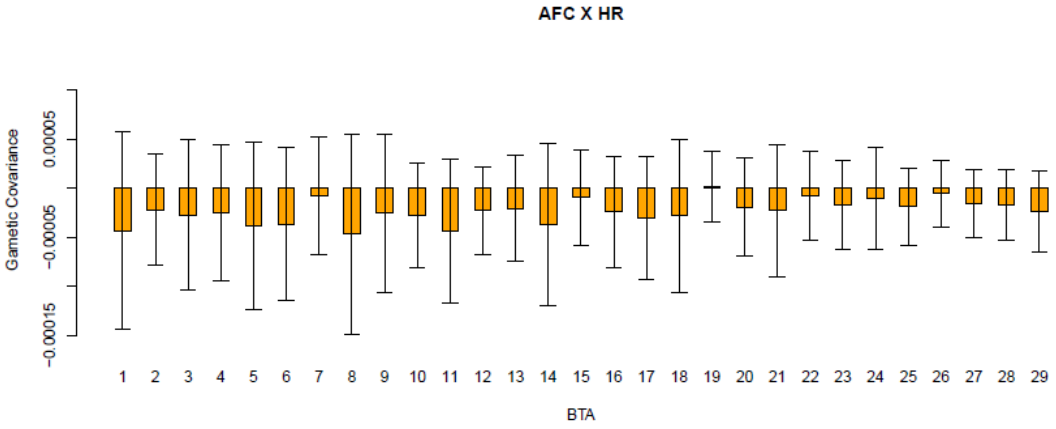
	Variance of Gametic Diversity				CRV			
Trait	Average	SD	Min	Max	Average	SD	Min	Max
AFC	1.62e-01	2.56e-02	7,73E-02	3,14E-01	0.46	0.03	0.34	0.61
HP	9.05e-05	1.40e-05	3,59E-05	1,70E-04	0.47	0.03	0.34	0.59
HR	6.10e-05	9.35e-06	2,26E-05	1,13E-04	0.47	0.03	0.34	0.59
Stay	1.49e-04	2.34e-05	7,24E-05	2,75E-04	0.46	0.03	0.33	0.58
SC	3.29e-03	4.98e-04	1,49E-03	6,47E-03	0.45	0.03	0.32	0.61
YW	5.45e-01	8.08e-02	2,79E-01	9,92E-01	0.46	0.03	0.33	0.58
WWd	2.91e-01	4.23e-02	1,35E-01	4,96E-01	0.45	0.03	0.31	0.56
WWm	2.08e-01	3.01e-02	1,03E-01	3,61E-01	0.46	0.03	0.32	0.57

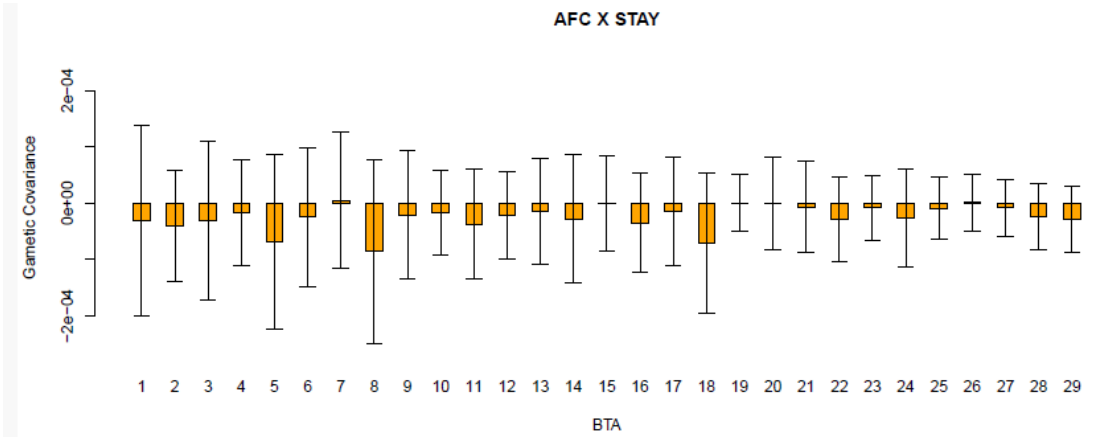
Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)

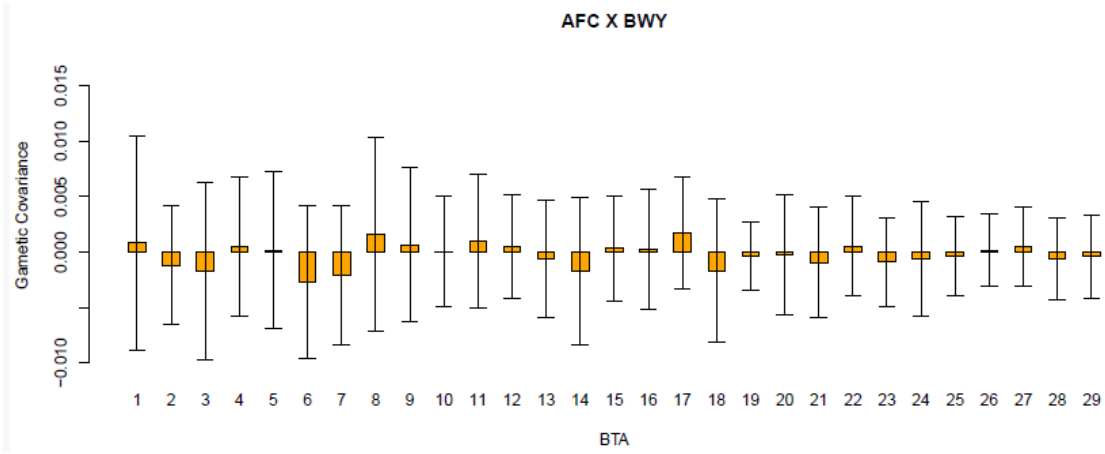
Table S2 - One-sample Kolmogorov-Smirnov test

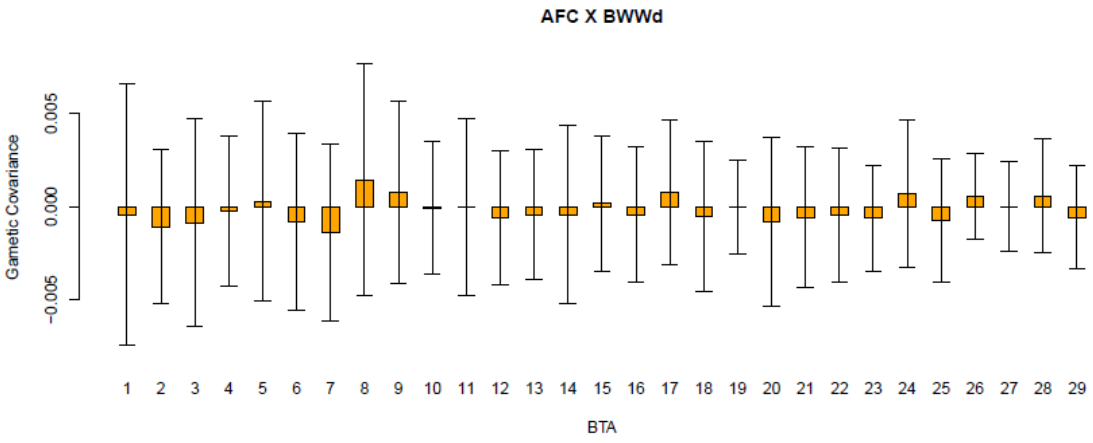
	Variance of Gametic Diversity		CRV	
Trait	D	p-value	D	p-value
AFC	0.028056	2.2e-16	0.0026573	0.7738
HP	0.029526	2.2e-16	0.0038357	0.3214
HR	0.028541	2.2e-16	0.0031178	0.5832
Stay	0.028637	2.2e-16	0.0033258	0.4992
SC	0.02738	2.2e-16	0.0035398	0.419
YW	0.029229	2.2e-16	0.0049081	0.1009
WWd	0.026761	2.2e-16	0.0035183	0.4267
WWm	0.99989	2.2e-16	0.0024935	0.8356

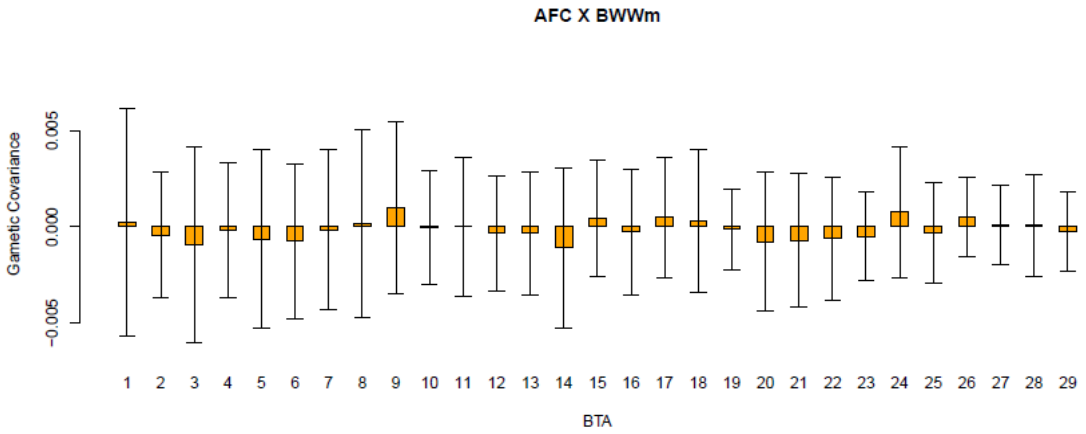
Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW); CRV: Coefficient of Relative Variation

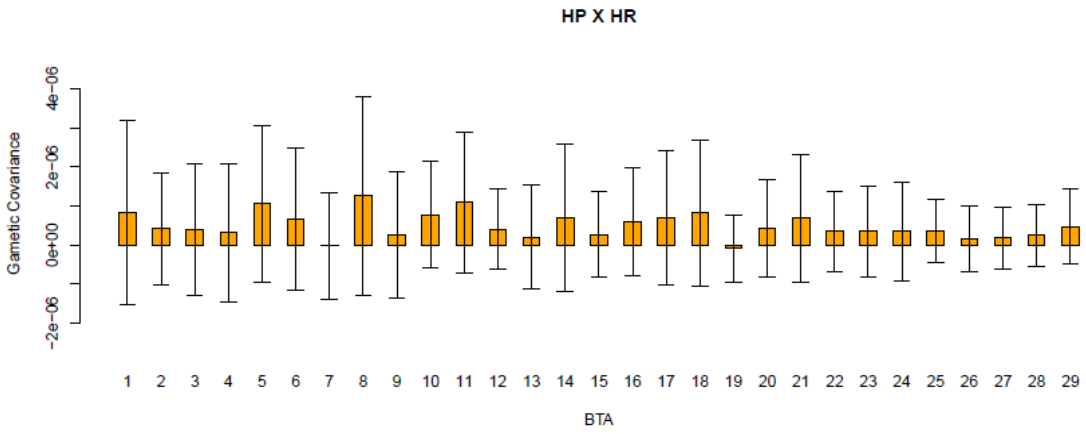

AFC X HR
Gametic Covariance
−0.00015
−0.00005
0.00005
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

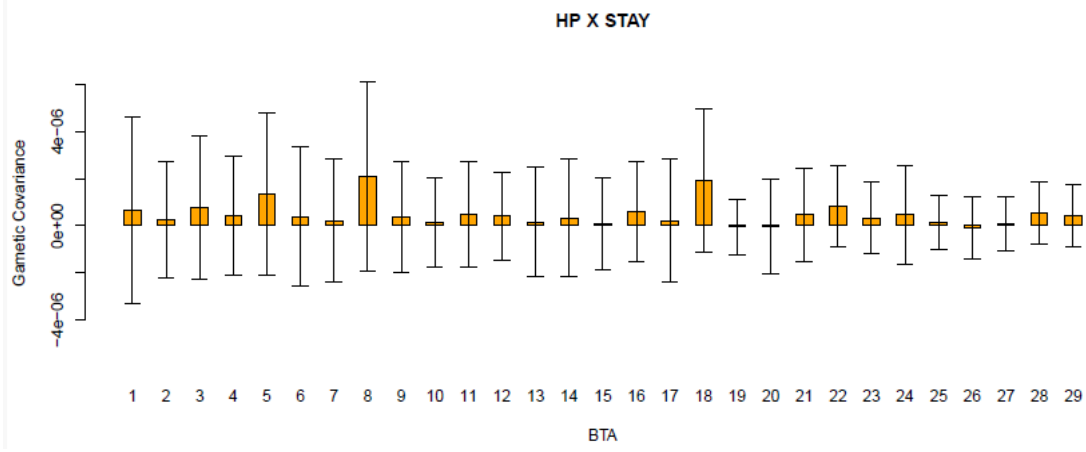

AFC X STAY
Gametic Covariance
−2e−04
0e+00
2e−04
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

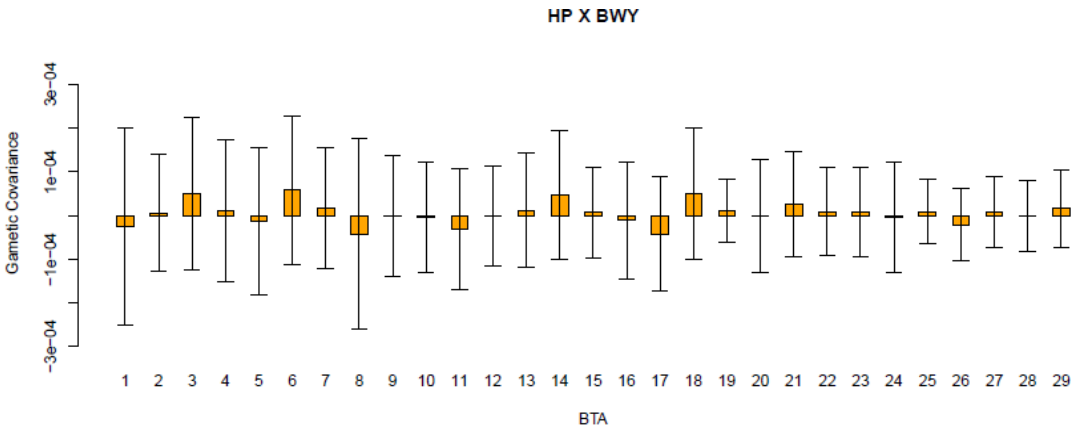

AFC X BWY
Gametic Covariance
−0.010
0.000
0.005
0.010
0.015
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

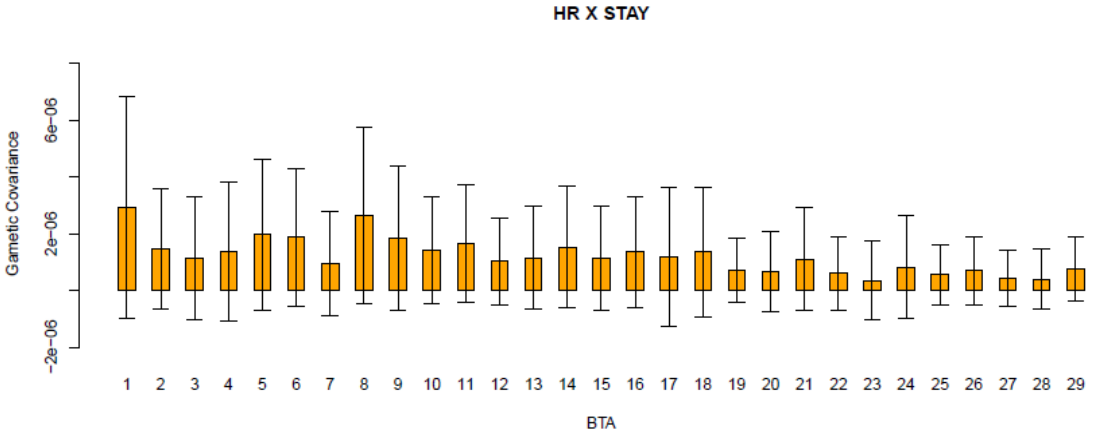

AFC X BWWd
Gametic Covariance
−0.005
0.000
0.005
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

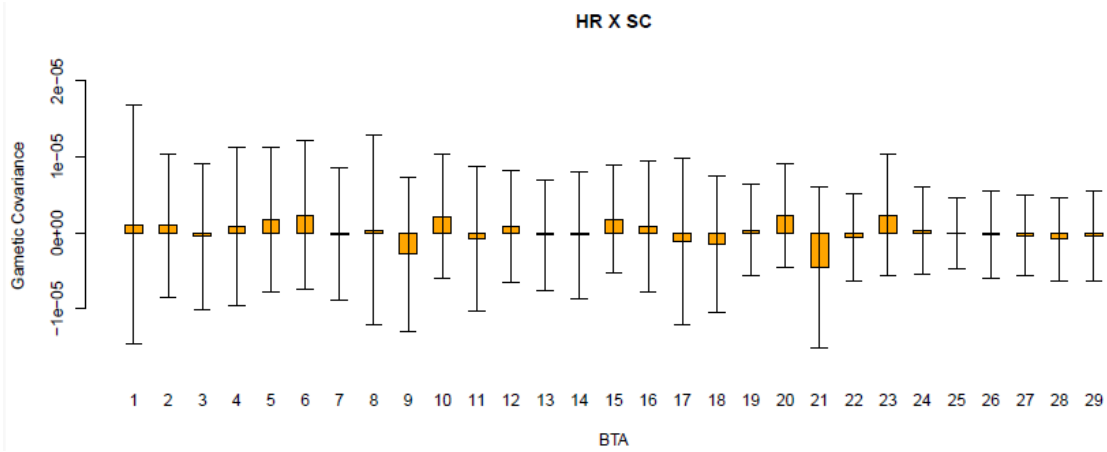

AFC X BWWm
Gametic Covariance
0.005
0.000
−0.005
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

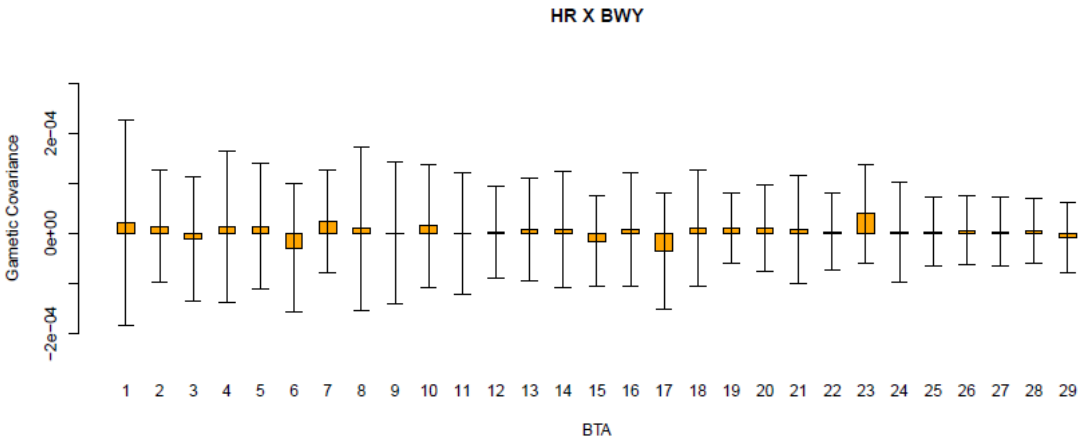

HP X HR
Gametic Covariance
4e−06
2e−06
0e+00
−2e−06
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

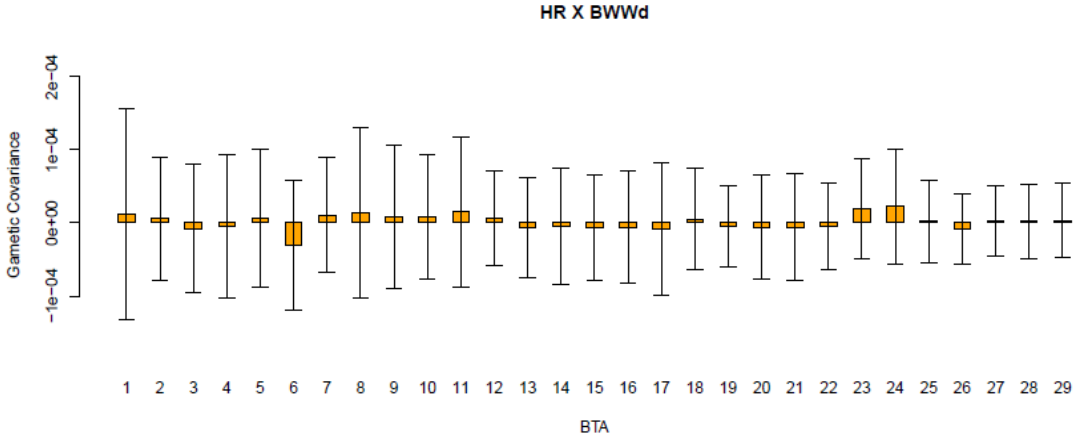

HP X STAY
Gametic Covariance
4e−06
0e+00
−4e−06
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

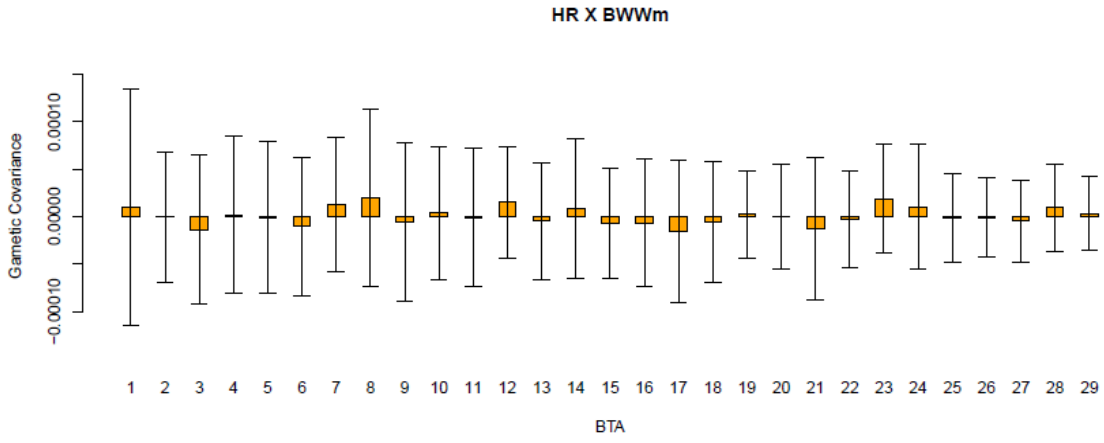

HP X BWY
Gametic Covariance
-3e-04
-1e-04
1e-04
3e-04
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

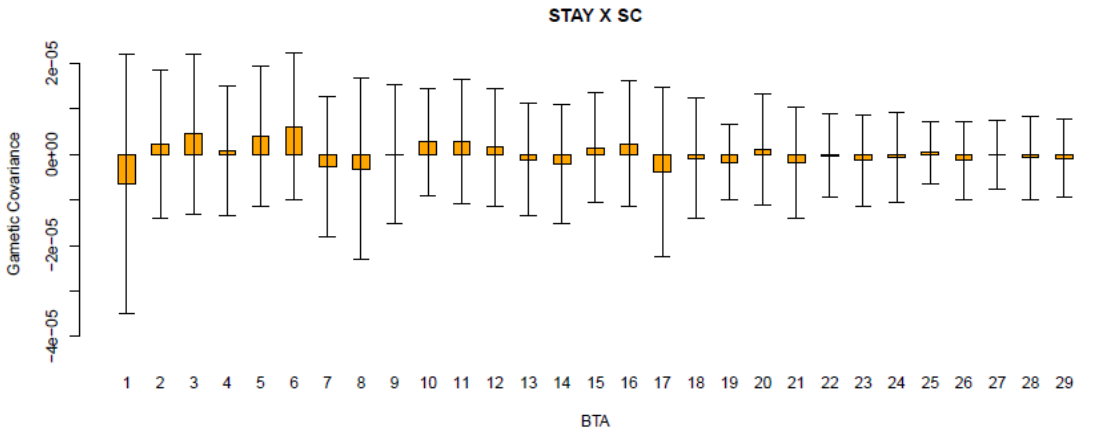

HR X STAY
Gametic Covariance
-2e-06
2e-06
6e-06
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

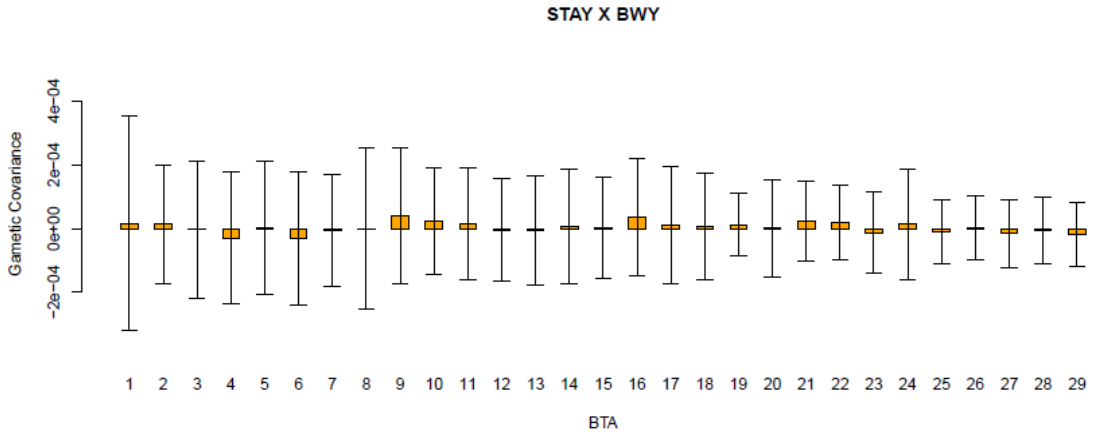

HR X SC
Gametic Covariance
-1e-05
0e+00
1e-05
2e-05
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

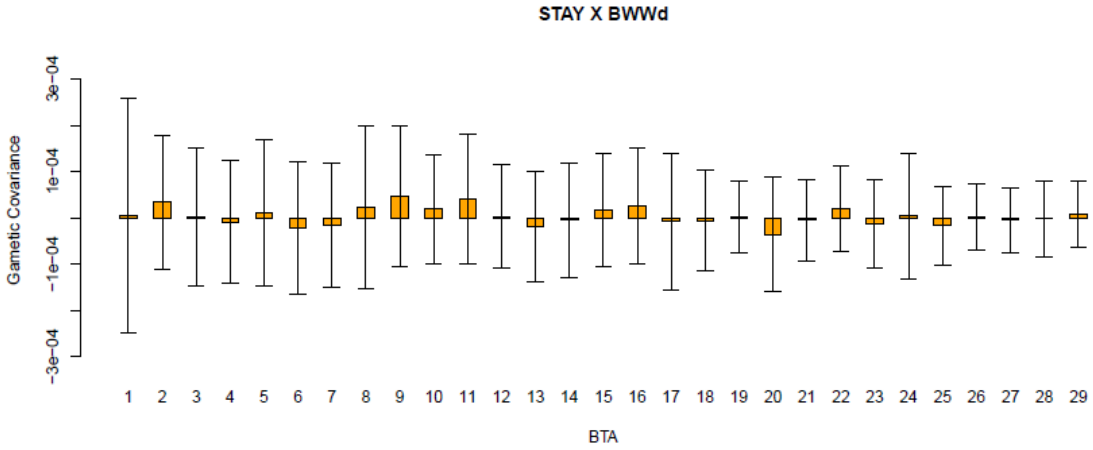

HR X BWY
Gametic Covariance
-2e-04
0e+00
2e-04
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

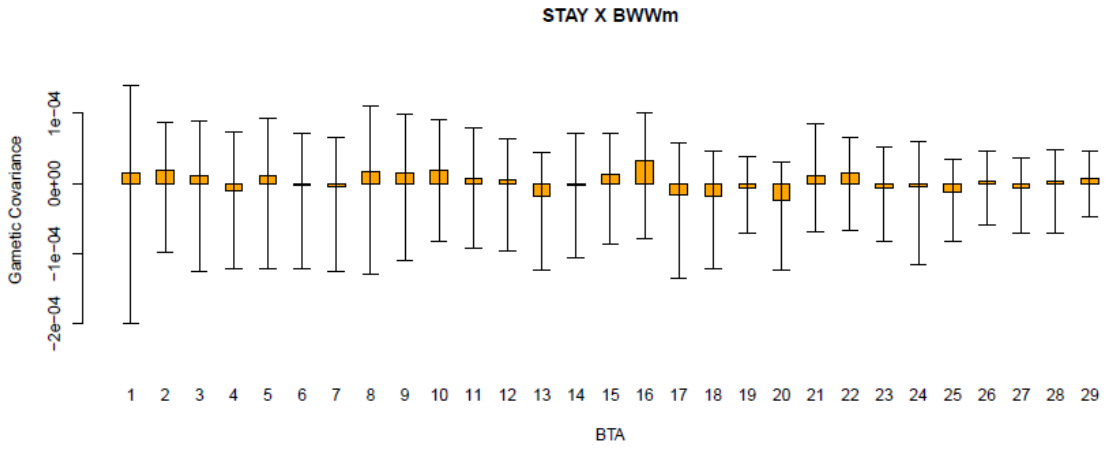

HR X BWWd
Gametic Covariance
-1e-04
0e+00
1e-04
2e-04
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

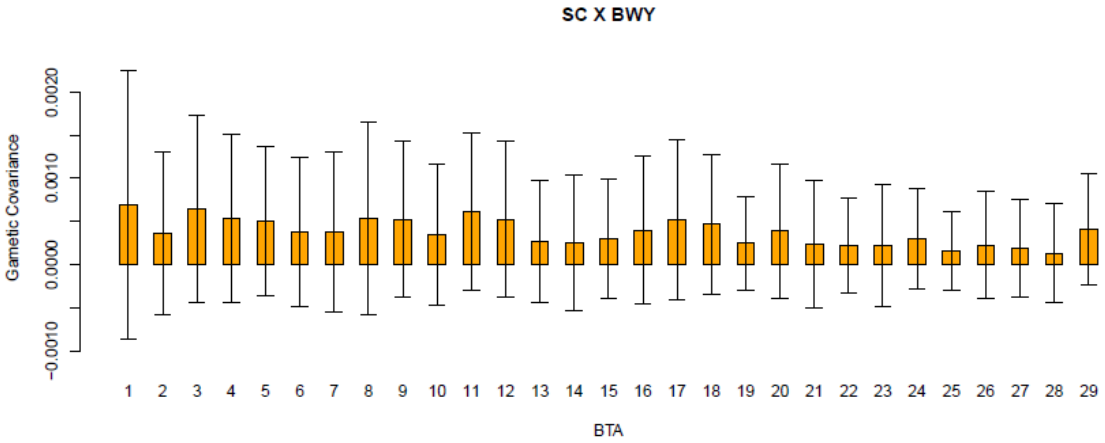

HR X BWWm
Gametic Covariance
-0.00010
0.00000
0.00010
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

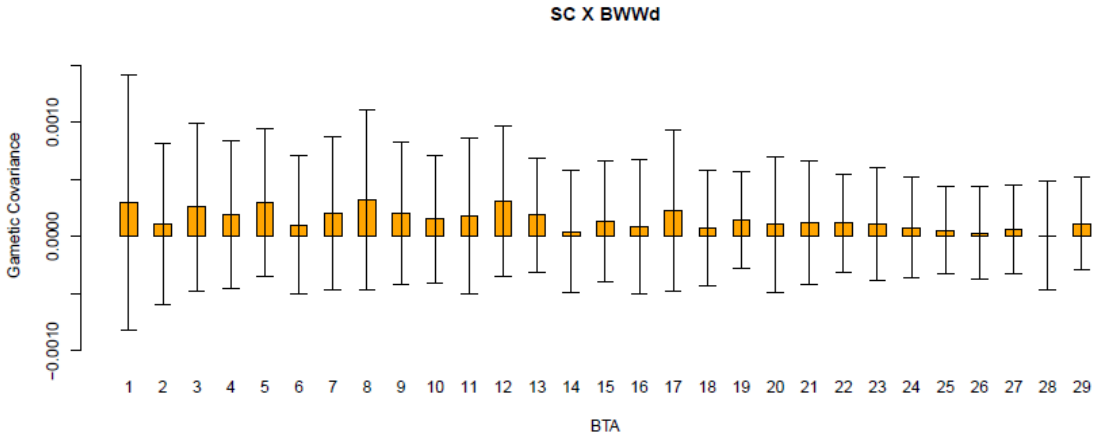

STAY X SC
Gametic Covariance
-4e-05
-2e-05
0e+00
2e-05
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA


STAY X BWY
Gametic Covariance
-2e-04
0e+00
2e-04
4e-04
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA


STAY X BWWd
Gametic Covariance
-3e-04
-1e-04
1e-04
3e-04
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA


STAY X BWWm
Gametic Covariance
-2e-04
-1e-04
0e+00
1e-04
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA


SC X BWY
Gametic Covariance
-0.0010
0.0000
0.0010
0.0020
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA


SC X BWWd
Gametic Covariance
-0.0010
0.0000
0.0010
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
BTA

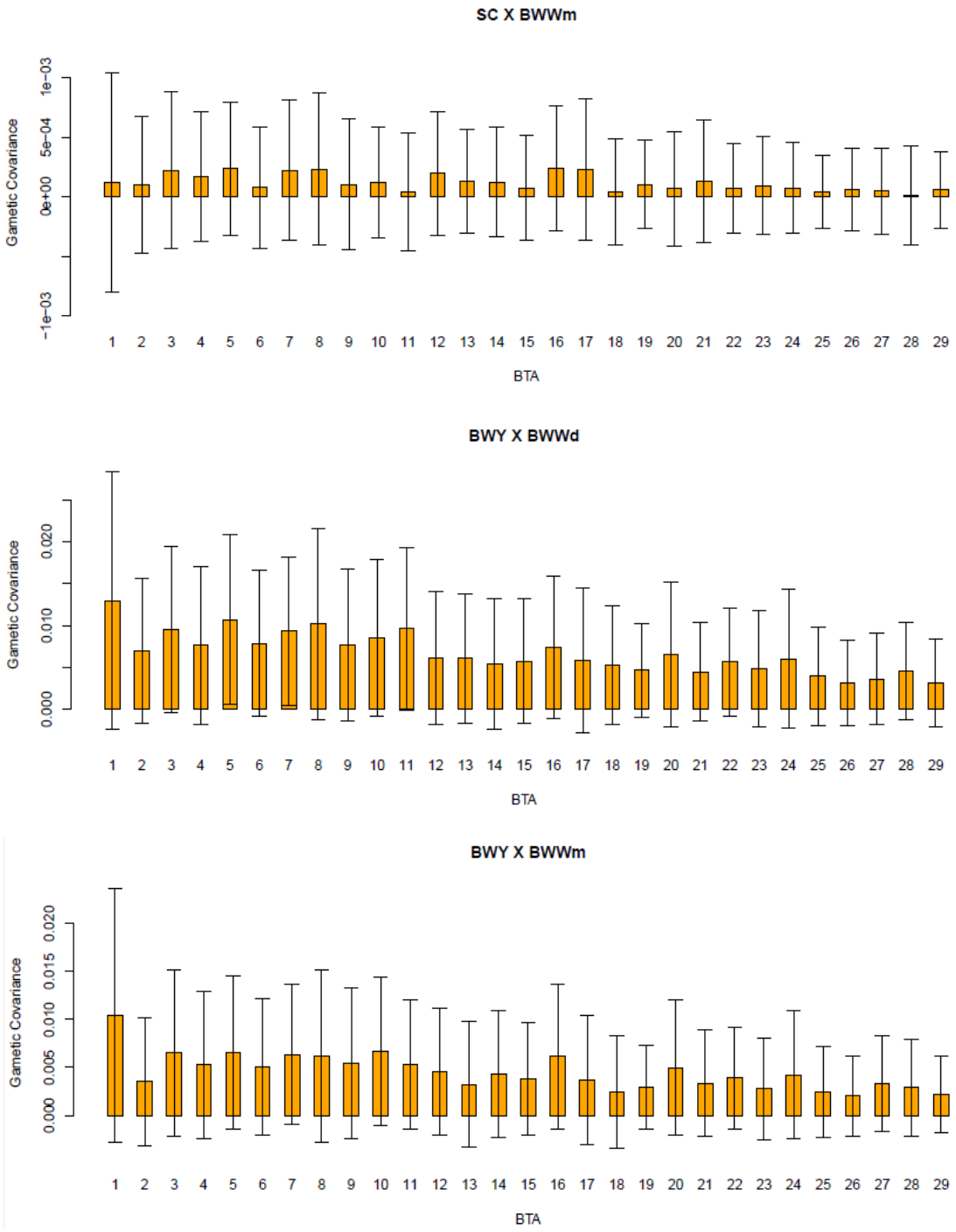


Figure S1 - Bar plots (distribution) of covariance of gametic diversity by chromosome between different traits. Bars indicate averages and whiskers represent standard deviations; Age at first calving (AFC); Heifer pregnancy (HP); Heifer rebreeding (HR); Stayability (STAY); Scrotal circumference (SC); Weaning weight direct (WWd); Weaning weight maternal (WWm) Yearling weight (YW)